

2018

Designing Sustainable Landscapes: Modeling Focal Species

Kevin McGarigal

University of Massachusetts Amherst

Brad Compton

University of Massachusetts Amherst

Ethan Plunkett

University of Massachusetts Amherst

Bill DeLuca

University of Massachusetts Amherst

Joanna Grand

University of Massachusetts Amherst

Follow this and additional works at: [https://scholarworks.umass.edu/
designing_sustainable_landscapes_techdocs](https://scholarworks.umass.edu/designing_sustainable_landscapes_techdocs)

McGarigal, Kevin; Compton, Brad; Plunkett, Ethan; DeLuca, Bill; and Grand, Joanna, "Designing Sustainable Landscapes: Modeling Focal Species" (2018). *Designing Sustainable Landscapes Project Technical Documents*. 9.

Retrieved from https://scholarworks.umass.edu/designing_sustainable_landscapes_techdocs/9

This Article is brought to you for free and open access by the The Designing Sustainable Landscapes (DSL) Project at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Designing Sustainable Landscapes Project Technical Documents by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Designing Sustainable Landscapes: Modeling Focal Species

A project of the University of Massachusetts Landscape Ecology Lab

Principals:

- Kevin McGarigal, Professor
- Brad Compton, Research Associate
- Ethan Plunkett, Research Associate
- Bill DeLuca, Research Associate
- Joanna Grand, Research Associate

With support from:

- North Atlantic Landscape Conservation Cooperative (US Fish and Wildlife Service, Northeast Region)
- Northeast Climate Science Center (USGS)
- University of Massachusetts, Amherst



Report date: 20 April 2018

Reference:

McGarigal K, DeLuca WV, Compton BW, Plunkett EB, and Grand J. 2017. Designing sustainable landscapes: modeling focal species. Report to the North Atlantic Conservation Cooperative, US Fish and Wildlife Service, Northeast Region.

Table of Contents

1	Problem Statement	3
2	Solution Statement	3
3	Detailed Description of Process.....	5
3.1	Representative Species.....	5
3.2	Climate Niche Model.....	6
3.2.1	CN model evaluation	13
3.3	Habitat Capability Model	14
3.3.1	Spatial data	17
3.3.2	Resistant kernels	20
3.4	Prevalence Model	23
3.5	Landscape Capability Model	23
3.5.1	LC model evaluation.....	26
3.6	Landscape Change Indices.....	31
4	Alternatives Considered and Rejected.....	37
5	Major Implementation Constraints.....	38
6	Major Risks and Dependencies	38
7	Literature Cited.....	39
	Appendix – Species Products	41

1 Problem Statement

To assess the capability of the landscape to support sustainable wildlife populations under various climate change and urban growth scenarios (as well as several other ecosystem drivers in subsequent project phases), reliable and informative species' climate niche and habitat capability models must be developed for a suite of representative species. A species-based approach to assessing the overall resiliency of the landscape to anthropogenic alterations, such as species' climate-habitat models, complements the coarse-filtered ecosystem-based assessment provided by the ecological integrity analysis as described in the ecological integrity document (McGarigal et al 2017), under the auspices of the Designing Sustainable Landscapes (DSL) project.

2 Solution Statement

We develop species' climate niche models, habitat capability models, and prevalence models and combine these into a single landscape capability model for a suite of representative species to assess the capability of the Northeast Region to sustain a suite of identified conservation priority species under future landscape change scenarios.

First, we use logistic regression methods to build species' *Climate Niche (CN)* models from downscaled climate data and independent species' occurrence data (e.g., Breeding Bird Survey) for the period 1985-2010. These models predict the probability of occurrence of each species based on their current geographic distribution in relation to several climate variables based on data representing the past 30 years. In the context of the *Landscape Change, Assessment and Design (LCAD)* model, we use these fitted models to predict the future distribution of the species' climate niche under alternative climate change scenarios. Importantly, we use these predictions to determine where the species might occur if they are able to immediately redistribute to remain within their current climate niche envelope (*CNE*), but they are not meant to predict where the species will actually occur because of our uncertainty in the species' ability to geographically track climate and the potentially limiting role of future habitat changes independent of climate, as well as time lags in habitat response to climate change. The result of the climate niche modeling is a map depicting the relative probability of occurrence (scaled 0-1) of each species at each time step under each climate change scenario based on the species' current climate niche. In addition, we convert the continuous relative probability of occurrence surface into a binary map representing the species' *CNE* by thresholding the probability surface to achieve the lowest commission (false positive) error rate (i.e., proportion of observed absences wrongly predicted to be present) given a specified model sensitivity of either 95%, 97.5% or 99% (i.e., percentage of observed occurrences correctly predicted to be present) based on empirical presence/absence data. Lastly, we summarize these results across climate change scenarios to reflect the uncertainty in future climate changes.

Second, we use the program HABIT@, a spatially explicit, GIS-based wildlife habitat modeling framework developed in the UMass Landscape Ecology Lab, to build species' habitat capability models. These models produce an index of habitat capability that we refer to as the *Habitat Capability (HC)* index for each species based on the condition of the landscape in relation to a suite of environmental variables. In the context of the LCAD

model, we use these HABIT@ models to predict the future habitat capability of the landscape under alternative land use (e.g., urban growth) scenarios. Importantly, we use these predictions to determine where the species might occur if they are able to immediately redistribute to track suitable habitat conditions, but they are not meant to predict where the species will actually occur because of our uncertainty in the species' ability to geographically track habitat changes and the potentially limiting role of future climate independent of habitat. The result of the HABIT@ modeling is a map depicting the *HC* index (scaled 0-1) for each species at each time step under each landscape change scenario. Lastly, we summarize these results across landscape change simulations to reflect the uncertainty in future landscape changes.

Third, we use kernel density estimators to build species' *Prevalence* models based on species' occurrence data (e.g., Breeding Bird Survey). These models predict the species' current distribution based solely on the species' observed spatial distribution independent of any explanatory variables and are intended to capture biogeographic factors influencing species' distributions that are not accounted for by the climate niche and habitat capability models. In the context of the LCAD model, we use these prevalence models to regulate the species' predicted probability of occurrence (see below) separately from that of climate and habitat. This is particularly important in some species' distributions where prevalence is less than would be expected based solely on climate suitability and habitat capability, presumably due to other biogeographic factors such as interspecific interactions and disease that we cannot measure directly. The result of the prevalence modeling is a map depicting the relative prevalence (0-1, although distributed as an integer grid scaled 0-100) of each species under the current landscape conditions.

Fourth, we synthesize the previous results for each species into a composite *Landscape Capability (LC)* index at each time step for each landscape change simulation. Specifically, we combine the species' *CN*, *HC* and *Prevalence* into a single index (*LC*) scaled 0-1 and use logistic regression to evaluate the predictive ability of the model based on independent species' occurrence data (e.g., eBird). Importantly, the *LC* models provide an index of species occurrence, not the true probability of occurrence. In the context of the LCAD model, we use the intersection of a species' *LC* map at any future timestep in relation to the initial or baseline condition in 2010 as the basis for summarizing the potential impacts of habitat and climate changes on a species. The result of the landscape capability modeling is a map depicting an index of occurrence for each species at each time step under each landscape change scenario.

Lastly, we assess the potential impacts of habitat and climate changes on each species using a variety of non-spatial and spatial indices. First, we compute a complementary set of non-spatial indices for each species based on the proportional change in *LC* due to climate change, habitat change, or both within the specified geographic extent. These non-spatial indices are primarily useful for establishing conservation objectives or targets for species in conservation design or for comparison among landscape change scenarios. The vulnerability indices are computed for each landscape change simulation based on the 2080 predictions and summarized as the mean across simulations.

Second, we derive a variety of spatial indices representing the species' potential response to climate change, habitat change or both based on changes in *LC* under different assumptions

or for different purposes. These spatial indices are useful for prioritizing locations for conservation action for each species in the context of landscape conservation design and for visualizing the potential changes in the distribution of a species due to climate change, habitat change or the combination of the two. These spatial indices are computed for each landscape change simulation based on the 2080 predictions only (for parsimony sake) and summarized as the mean across simulations, and to facilitate their use in landscape conservation design each of these spatial indices is quantile-scaled within the project area.

3 Detailed Description of Process

Our species-based approach for evaluating the Region's capability to sustain priority wildlife species under future landscape change scenarios involves six major steps: 1) selecting a suite of representative species, 2) developing a climate-niche model for each species, 3) developing a habitat capability model for each species, 4) developing a prevalence model for each species, and 5) combining the results of the climate, habitat, and prevalence models into a landscape capability model for each species at each time step under each landscape change scenario to quantify uncertainty in the predictions of species occurrence, and 6) computing the non-spatial and spatial landscape change indices for each species. Each of these steps are described in this section. However, this section is intended to provide a general overview of the process; a description the landscape capability model for each focal species is provided in separate documents (**Appendix A**).

3.1 Representative Species

The first step involves selecting species for modeling. Importantly, we developed a modeling framework for assessing landscape capability for any focal species regardless of the purpose of the selected species. For example, individual species models can be developed for representative species, indicator species, threatened and endangered species, vulnerable species, flagship species, game species or any other species of conservation interest. Currently, we are focusing on developing models for a suite of representative species under the assumption that these relatively few species serve as surrogates for the much large suite of conservation priority species.

In a separate project led by UMass and directed by the U.S. Fish and Wildlife Service (USFWS), we selected a total of 87 terrestrial species to represent clusters of similar terrestrial ecological systems within the NALCC (NALCC 2011). For the first phase of this project, we selected a subset of 10 of these representative species for landscape capability modeling. For this purpose, a tiered approach was developed by Scott Schwenk at the University of Vermont (currently with USFWS and director of science for the NALCC) to prioritize species. Species listed in Tier 1 received the highest priority during phase 1. An additional 20 species were selected from this tiered list for modeling in the second phase of this project. Species assignments to tiers were based on both ecological and feasibility criteria. In general, Tier 1 species: 1) represent habitat clusters that account for large portions of the NALCC; 2) are sensitive to anthropogenic disturbances; 3) are relatively well understood; and 4) are cases for which the required environmental data are readily obtainable. The final suite of species modeled in phase 1 and phase 2 are listed in **Table 1**.

Table 1. List of 30 representative species modeled for the Northeast region. B=breeding season; NB=nonbreeding season; A=all year; M=migratory.

Species	Species
American black duck (B)	Marsh wren (B)
American black duck (NB)	Moose (A)
American oystercatcher (B)	Northern Waterthrush (B)
American woodcock (B)	Ovenbird (B)
Bicknell's thrush (B)	Prairie warbler (B)
Black bear (A)	Red-shouldered hawk (B)
Blackburnian warbler (B)	Ruffed grouse (B)
Blackpoll warbler (B)	Saltmarsh sparrow (B)
Box turtle (A)	Sanderling (M)
Brown-headed nuthatch (B)	Snowshoe hare (A)
Cerulean warbler (B)	Snowy egret (B)
Common loon (B)	Virginia rail (B)
Diamondback terrapin (A)	Wood duck (B)
Eastern meadowlark (B)	Wood thrush (B)
Louisiana Waterthrush (B)	Wood turtle (A)

It is important to remember that each representative species selected for modeling is presumed to function as a representative for a much larger suite of priority species associated with a cluster of ecological systems. The species assemblages and ecological system clusters represented by each species are documented separately (see http://www.fws.gov/northeast/science/representative_species.html).

3.2 Climate Niche Model

The next step is to develop a *Climate Niche (CN)* model for each selected representative species. Briefly, a *CN* model represents a species' apparent range of tolerance to climate variability based on the species' current distribution and the current climate. In the context of the LCAD model, we use the species' *CN* to predict the future distribution of the species' climate niche under alternative climate change scenarios. Importantly, these predictions are used to determine where suitable climate conditions will exist based on current climate associations, not where they will actually occur, since the latter depends on whether they can and choose to get there from currently occupied sites. A general description of the *CN* modeling approach follows and is illustrated using results for wood thrush. The process

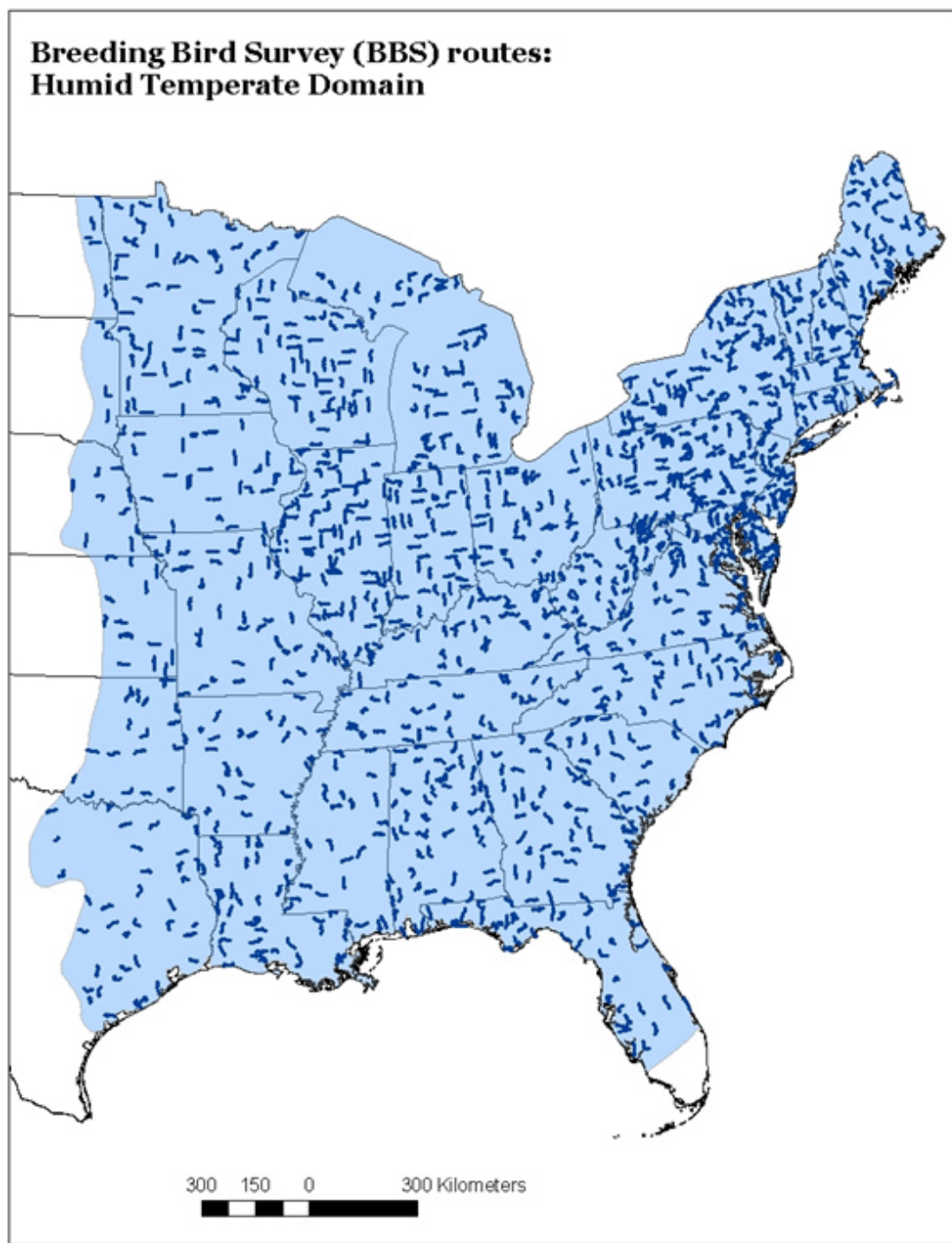


Figure 1. Geographic extent of the assessment area used to develop and apply species' climate niche models. Specifically, empirical observations of species' presence and absence are limited to Bailey's Humid Temperate Domain within the US (shaded blue area). The lines shown here depict the Breeding Bird Survey routes within this extent that are potentially available for inclusion in species' climate niche models.

described next is the same for all species, even though some of the details, such as sources of empirical data, vary slightly among species.

First, we define the geographic extent for compiling climate and species' occurrence data as the Humid Temperate Domain (sensu Bailey et al. 1994) within the U.S. (**Fig. 1**). We restricted our geographic extent to the U.S. because most of the extant data used for model building (e.g., Breeding Bird Survey, eBird, Mountain Birdwatch) is limited to the U.S. We limited the geographic extent to the Humid Temperate Domain in order to limit the distance between absent and present data in geographic and climatic space which, if ignored, can lead to high rates of model commission errors (Lobo et al. 2010). By doing so, we avoid situations where relatively similar but spatially distant climate conditions are included in models as both suitable (based on presence locations) and unsuitable (based on absence locations). For example, wood thrush are abundant in the seasonable, wet climate of the Great Smokey Mountains. A model restricted to the Humid Temperate Domain would likely capture this signal. However, a model that considered the entire U.S. would identify the seasonable, wet conditions of the Pacific Northwest (well outside of the range of wood thrush) as unsuitable (due to lack of occurrences) and thus the wood thrush's propensity for seasonable, wet climate conditions in the eastern U.S. might not be identified by the model. Conversely, the geographic extent must also provide enough geographic range to capture locations where individuals are absent due to unfavourable climatic conditions, as opposed to absences solely due to ecological constraints, such as competition or dispersal limitations (Lobo et al. 2010). Lastly, the geographic extent must also include climate variation beyond that found in the project area (Northeast) to account for future climate changes within the project area. In particular, the geographic extent must extend south and west far enough to include current climate conditions that are likely to be found within the project area over the next 70 years. Given all of these considerations, we consider the Humid Temperate Domain within the U.S. to be a reasonable compromise.

Second, we identify a suite of climate variables that we deem, *a priori*, important in defining a species' climate niche based on a review of the literature (e.g., Lawler et al. 2009, Wiens et al. 2009, Matthews et al. 2011) and availability of data (see technical document on climate, McGarigal et al 2017, for details) . For the species considered thus far, these climate variables include:

- 1) Average annual precipitation (precip)
- 2) Growing season precipitation (precipgs)
- 3) Average annual temperature (temp)
- 4) Average minimum winter temperature (tmin)
- 5) Average maximum summer temperature (tmax)
- 6) Growing degree days (gdd)
- 7) Heat index 35 (heat35)

Values for 2010 are derived from the Parameter-elevation Relationships on Independent Slopes Model (PRISM) dataset developed by Oregon State University with support from the U.S. Department of Agriculture through the Natural Resources Conservation Service (USDA-NRCS). This model uses a weighted climate-elevation regression approach to model the temperature and precipitation in each digital elevation model (DEM) grid cell. To develop the regression model in each cell, the model considers the most similar of 10,000

and 13,000 stations (for temperature and precipitation, respectively) in physiographic space, including the factors: location, elevation, coastal proximity, aspect, vertical atmospheric layer, topographic position, and orographic effects. The PRISM data are available as 30-year normal grids of the entire U.S. consisting of 800m cells with monthly average precipitation and monthly average minimum and maximum temperatures averaged across the years 1985 – 2010. Given these baseline values, we develop future projections of these climate variables based on the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 5 (CMIP5) multi-model dataset. Briefly, we process the output from 14 *Atmospheric-Ocean General Circulation Models* (AOGCMs) downscaled to 12 km using the *Bias Corrected Spatial Disaggregation* (BCSD) downscaling approach to create an ensemble average AOGCM projection under each of 2 Representative Concentration Pathways (RCP) scenarios (RCP4.5 and RCP8.5) (see technical document on climate, McGarigal et al 2017, for details).

Third, we compile an independent data set on species occurrence within the assessment area. Our goal is to consistently use Breeding Bird Survey (BBS) data (<https://www.pwrc.usgs.gov/bbs/>) to develop climate niche models for bird species. However, in the event that BBS data does not provide adequate samples, for example with non-avian species, habitats that are not adequately sampled (e.g., salt marsh) or avian species with very restricted distributions, we augment this dataset with other existing data as needed and deemed appropriate (e.g., Mountain Birdwatch, MBW, data). Note, for some bird species BBS data alone or in combination with these supplementary datasets was not adequate to develop CN models. For these species, we use the eBird data described below under model evaluation to develop the CN model. Unfortunately this precludes eBird as a source of data for evaluation for these species, as discussed below.

BBS data is maintained by the US Geological Survey's Patuxent Wildlife Research Center. For suitable species, we summarize the BBS data as follows:

- 1) We use data from 1990 – 2010, as this time period matches the PRISM data and is likely to reflect the current range of climatic conditions a species occupies.
- 2) Based on the recorded detections, we consider the presence/absence of each species within each survey route segment (~8 km road length containing 10 point counts), not the total count. We compute the proportion of years in which each species was present on each segment. Note, because BBS does not maintain segment-level spatial data, we divide each 40 km route into five equal segments and use the starting location to order each segment.

Fourth, we use multiple logistic regression to predict the response as a function of the suite of climate variables, with the following modeling considerations: 1) we treat each BBS route segment as an independent observation and make no attempt to account for residual spatial autocorrelation as it is not deemed relevant given our modeling purpose, 2) we assume a binomial error distribution for the proportional response and treat the number of survey years at a segment as the trial size, 3) we do not account for imperfect detection of species in the model as it is not feasible given the sampling design (without violating important assumptions) and is not deemed essential given our modeling purpose, and 4) we consider quadratic terms for each of the predictors to account for nonlinear

relationships (in the logit of the response). We conduct an all subsets logistic regression on the climate predictors that evaluates every combination of climate predictors (in linear and quadratic form). To select the best model, we consider models that minimize prediction errors of commission (false positives) given a specified model sensitivity of 95%, 97.5% and 99% (depending on species). Model sensitivity is the proportion of "present" observations that are correctly predicted to be "present" (the complement of sensitivity is the omission error rate). We consider sensitivity levels of 95-99% because we want a prediction envelope that captures most of the species' observed distribution (i.e., a "present" envelope), but one that is not influenced by the most extreme observations. Given a particular model sensitivity, we seek to minimize the commission error (or false positive) rate; i.e., the model that results in the fewest "absent" observations incorrectly predicted to be "present". Note, the regression model predicts the relative probability of occurrence for each observation; a cutpoint or threshold value on the response scale (i.e., probability scale) must be selected in order to classify observations as "present" or "absent". In our case, we select the model with the lowest AIC value (goodness-of-fit criterion) within a delta false positive rate of 0.1 (i.e., within 10% of the best model's false positive rate). We consider this approach as a method to simultaneously minimize errors of commission while maintaining model parsimony, ensuring that of the models that do the best job of minimizing false positive rates, we select the most parsimonious possibility. This model selection procedure is similar to the conventional approach based on minimizing a goodness-of-fit criterion such as AIC, but here our goodness-of-fit criterion is first based on commission error given a specified model sensitivity. Lastly, we consider the range extent of the *CNE* derived from the "best" model at each sensitivity level (95%, 97.5%, and 99%) and select the model (or sensitivity level) that defines a *CNE* that best approximates the species' current geographic range and extent as defined by various sources such as breeding bird atlases (<http://www.pwrc.usgs.gov/bba/index.cfm?fa=bba.MapViewer>), eBird data, and the species published range (Poole 2008).

The approach described above results in a fitted logistic regression model (i.e., parameter estimates), which we use to define the species' *CN*, and a cutpoint on the response scale, which we use to define the species' *CNE*. Based on this model, we generate the following spatial results:

- 1) *Climate niche (speciesCN)* — map depicting each species' relative probability of climate suitability based on the climate in 2010 (**Fig. 2a**). Note, these maps depict climate suitability as a continuous surface scaled 0-1. A binary map depicting areas "inside" versus "outside" each species' *CNE* based on the climate in 2010 is not distributed but can be derived by thresholding the continuous probability surface at the specified cutpoint (e.g., **Fig. 2b**).
- 2) *Future climate niche (speciesCN2080)* — map depicting each species' relative probability of climate suitability based on the predicted climate in 2080 under the RPC 4.5 and 8.5 climate change scenarios. Note, these maps depict climate suitability as a continuous surface scaled 0-1 and are averaged across climate scenarios for the distributed product. A binary map depicting areas "inside" versus "outside" each species' *CNE* based on the predicted climate in 2080 averaged across the RPC 4.5 and

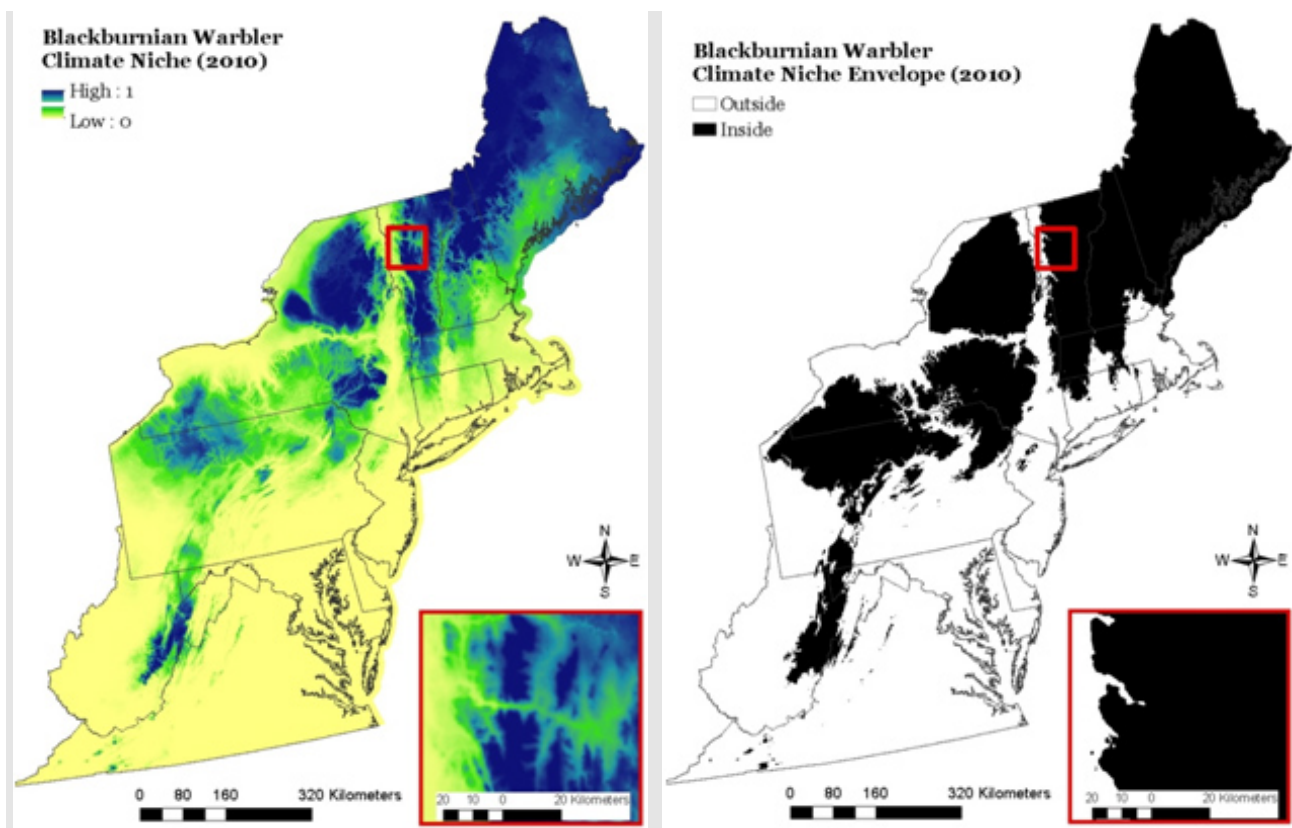


Figure 2. Example of a species' Climate Niche (CN) index, expressed as a continuous relative probability of climate suitability surface, and corresponding Climate Niche Envelope (CNE), derived by thresholding the CN index at the cutpoint that minimizes the commission error rate (false positives) given a specificity (true positives) rate of 95-99%, expressed as a binary map of where the species' is most likely to occur (black) based on climate suitability; shown here for the blackburnian warbler in the Northeast in 2010.

8.5 climate change scenarios is not distributed but can be derived by thresholding the continuous probability surface at the specified cutpoint.

- 3) *Climate zones (speciesCZ2080)* — maps derived by intersecting a species' current and future CNE (averaged across RPC 4.5 and 8.5 climate change scenarios) to delineate three distinct zones of uncertainty in the predicted future distribution of a species based on climate suitability: 1) *Zone of Persistence* - overlap of the current and future CNE; thus where we have high confidence in the species' predicted future occurrence; 2) *Zone of Contraction* - current CNE outside of the future CNE; where the future climate is no longer suitable and thus where we have lower confidence in the species' predicted future occurrence due to unknown population time lags and other factors; 3) *Zone of Expansion* - future CNE outside of the current CNE; where the future climate becomes suitable but is not currently suitable and thus where we have lower confidence in the species' predicted future occurrence due to unknown population time lags and other factors (**Fig. 3**). Note, these climate zones are used to derive some of the landscape change indices (described below).

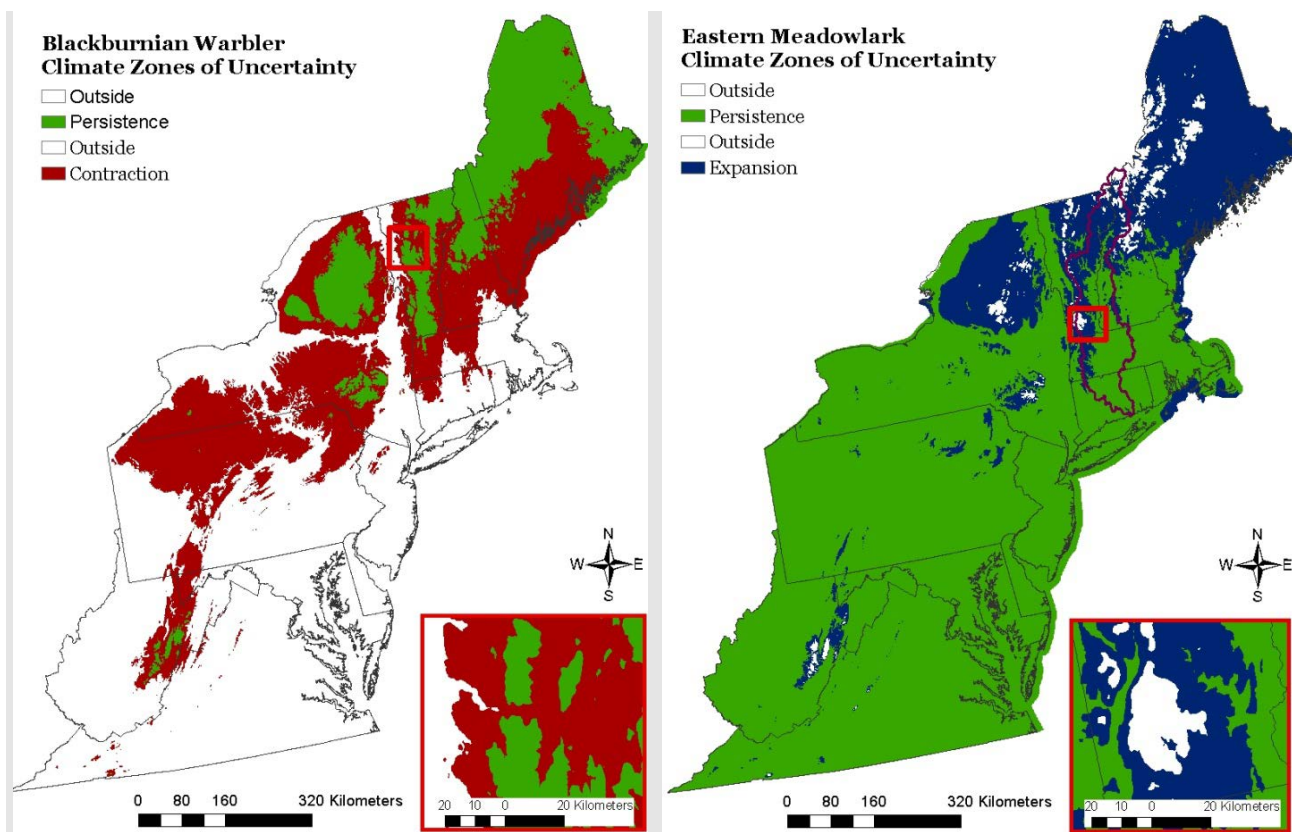


Figure 3. Climate Zones in a species' response to climate change, derived by intersecting the species' current Climate Niche Envelope(CNE) and mean future CNE (averaged across climate change scenarios in 2080), including: 1) zone of persistence, where the current and future CNE overlap, 2) zone of contraction, where the current CNE does not overlap the future CNE, and 3) zone of expansion, where the future CNE does not overlap the current CNE. Shown here for blackburnian warbler and eastern meadowlark in the Northeast.

In addition, we generate the following non-spatial results:

- 1) *Climate niche (ha)* — sum of *CN* in 2010 multiplied by 0.09 to convert it to maximum *CN*-equivalent hectares; it ranges from 0 (no suitable climate) to the study area extent in hectares (when the entire study area is all optimal climate).
- 2) *Percent change in climate niche* — sum of future *CN* in 2030 or 2080 (averaged across the RCP 4.5 and 8.5 climate change scenarios) minus the sum of *CN* in 2010, divided by the sum of *CN* in 2010, multiplied by 100 to convert to a percentage; it ranges from -100 with the complete loss of suitable climate and is unbounded on the upper end. Values <0 reflect shrinking *CN* whereas values >0 reflect expanding *CN* within the project area (**Table 2**).

Table 2. Example species' Climate Niche (CN) index in the Northeast region in 2010 and Climate Change indices for 2030 and 2080. The CN index is computed as the sum of the CN values across the landscape, reported in hectares, where the CN values represent the probability of suitable climate (see text for details) based on a model built from the species' known distribution in 2010. Statistics reported for 2030 and 2080 represent the percentage change relative to the CN index in 2010 averaged across two climate scenarios (RCP 4.5 and 8.5); 0 represents no change in CN.

Species	Climate Niche (ha)	Percent change in Climate Niche	
	2010	2030	2080
American woodcock	1,132,751,220	-1.6	-8.2
Blackburnian warbler	914,694,418	-24.4	-71.3
Blackpoll warbler	123,144,883	-48.6	-87.6
Eastern meadowlark	664,094,840	22.3	68.2
Louisiana waterthrush	231,888,468	25.0	102.3
Marsh wren	807,983,242	13.8	62.1
Moose	1,254,750,734	-9.2	-10.4
Northern waterthrush	707,764,917	-30.3	-68.4
Prairie warbler	163,586,333	30.4	122.6
Ruffed grouse	1,199,451,492	-15.4	-62.1
Wood duck	432,906,109	10.6	30.6
Wood thrush	1,458,459,115	-0.1	-0.1
Wood turtle	1,126,357,571	6.7	7.6

3.2.1 CN model evaluation

To assess *CN* model performance, we conduct a Monte Carlo randomization test of the model commission error rate. Specifically, for the final logistic regression model, we compute the observed commission error rate (i.e., false positive error rate) and compare it to the distribution of commission error rates under the null distribution. The null distribution represents the distribution of commission error rates expected if the species' occurrences are distributed randomly with respect to climate (i.e., climate has no real relationship to their distribution). We generate the null distribution by randomly shuffling the "present" and "absent" among the sample locations and then recomputing the commission error rate given the corresponding sensitivity level (95%, 97.5% or 99%) selected for the species, and repeating this process 1,000 times to generate a distribution of values. We expect the original model to have a smaller commission error rate than the bulk

of the null distribution if the species' distribution is nonrandom with respect to climate, therefore we compute the lower (one-sided) p -value as the proportion of the null distribution less than or equal to the observed commission error rate (**Fig. 4**). We interpret a significant difference (i.e., p -value ≤ 0.05) as confirmation of the reliability of the *CN* model.

3.3 Habitat Capability Model

After describing the key habitat requirements for a representative species (based on extensive literature review and consultations with experts), we model the quantity, quality and accessibility (i.e., capability) of habitat across the landscape at each time step under each landscape change simulation with HABIT@. Note, we use the term "habitat capability" instead of "habitat suitability" to distinguish our approach from conventional *Habitat Suitability Index (HSI)* modeling. In particular, our approach involves assessing not only the quantity and quality of habitat, typical of conventional habitat suitability models, but also the accessibility of quality habitat, which is unique to our approach. In the context of the LCAD model, we use HABIT@ to predict the future distribution of the species' habitat capability under alternative climate change and land use scenarios. Importantly, we use these predictions to determine where the species might occur based on their habitat requirements, not where they will actually occur, since the latter will depend on whether they can and choose to get there from currently occupied sites. A general description of the HABIT@ modeling framework follows; details of the HABIT@ models for each species are included in **Appendix A**.

HABIT@ is a multi-scale GIS-based system for modeling wildlife habitat capability developed by Brad Compton and Kevin McGarigal in the UMass Landscape Ecology Lab. *Habitat capability* refers to the ability of the environment to provide the local resources (e.g., food, cover, nest sites) needed for survival and reproduction in sufficient quantity, quality and accessibility to meet the life history requirements of individuals and local populations. Rather than being focused on a particular species or ecosystem, HABIT@ is intended to be general enough to model any animal (or perhaps plant) species. Species

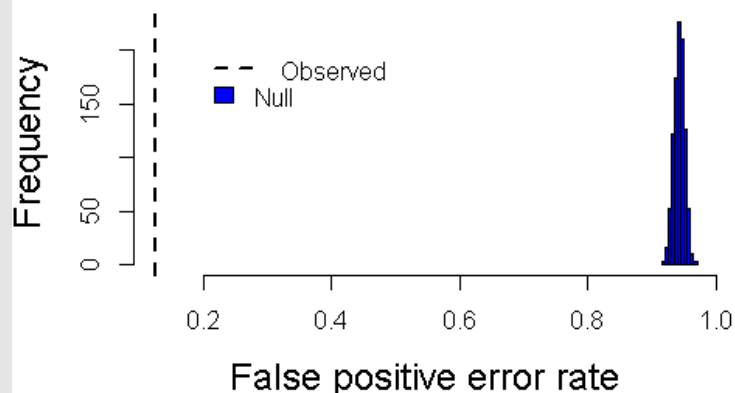


Figure 4. Monte Carlo randomization test (1,000 replications) of the observed commission error rate, given a model sensitivity of 95%, for the logistic regression model predicting species' presence from a suite of climate variables that minimized commission error at the specified model sensitivity; shown here for the wood thrush based on Breeding Bird Survey. The null distribution of commission error rates is given by the frequency distribution, while the observed commission error is shown as a vertical dashed line. The p -value (<0.001) is computed as the proportion of the null distribution less than or equal to the observed commission error. Shown here for the Blackburnian warbler in the Northeast.

models are typically developed directly by biologists with the best understanding of the species being modeled, without a lot of GIS or computer experience, although technical assistance is generally required to implement the models. HABIT@ was developed using the programming language APL, but is currently implemented using the HABIT@ scripting language.

HABIT@ models are static — HABIT@ does not model population dynamics nor population viability. Like an HSI model, the results are relative measures of habitat capability—they do not necessarily correspond to animal density or fitness. HABIT@ is not an individual-based model, and it does not explicitly model animal movement (although movement can be accounted for implicitly). HABIT@ models are, of course, limited by the availability, scale, and accuracy of available data, and the applicability of these data to the species being modeled. As with all habitat modeling, the greatest limitation is usually our lack of knowledge of the habitat requirements of the species being modeled. HABIT@ models are only as good as the biological information used to build them.

HABIT@ is spatially-explicit — the habitat capability value at each cell is dependent not only on the resources available at that cell, but on resources available in the neighborhood (e.g., is there enough forage to support an individual's homerange?), on the configuration of resources (e.g., are they juxtaposed or contiguous?), on impediments to movement (e.g., are food and nesting resources across an impassable road from each other?), and/or on the density of roads or development in the neighborhood. The models can be as complex as the biological understanding warrants and the available GIS data permit.

HABIT@ models are based on GIS grids representing environmental variables such as cover type, stand age, canopy density, slope, hydrological regime, as well as roads and development. Input grids can represent anything pertinent to the species being modeled, depending only upon the availability of data. Complex derived grids representing specialized environmental variables (such as stream channel constraints, cliffs suitable for nesting, or rainfall patterns) can be incorporated in HABIT@ models as well. Grids can be at any scale appropriate to the species being modeled, subject to data availability. GIS data that are available in a vector format can easily be converted to grids at an appropriate scale.

HABIT@ models consider habitat at two levels of organization (**Fig. 5**):

- 1) *Local Resource Capability (LRC)* — the availability of resources important to the species' life history, such as food, cover, or nesting, at the local, finest scale (a single grid cell or pixel); resources include factors influencing survival (e.g., food, thermal protection) and/or reproduction (e.g., nest substrate, protection from predators). Importantly, local resource availability is a function of both the ecological composition of the focal cell (i.e., cover type) and its spatial context (e.g., proximity to landscape features), but it is fundamentally an attribute of a focal cell and not a characterization of the landscape (or local neighborhood) surrounding the cell. A HABIT@ model consists of one or more local resources (e.g., cover, food, nesting) and the availability of each local resource is measured using one or more local resource indices that essentially translate environmental data (e.g., land cover) into 0-1 indices that can be combined into a composite *LRC* index for each resource.

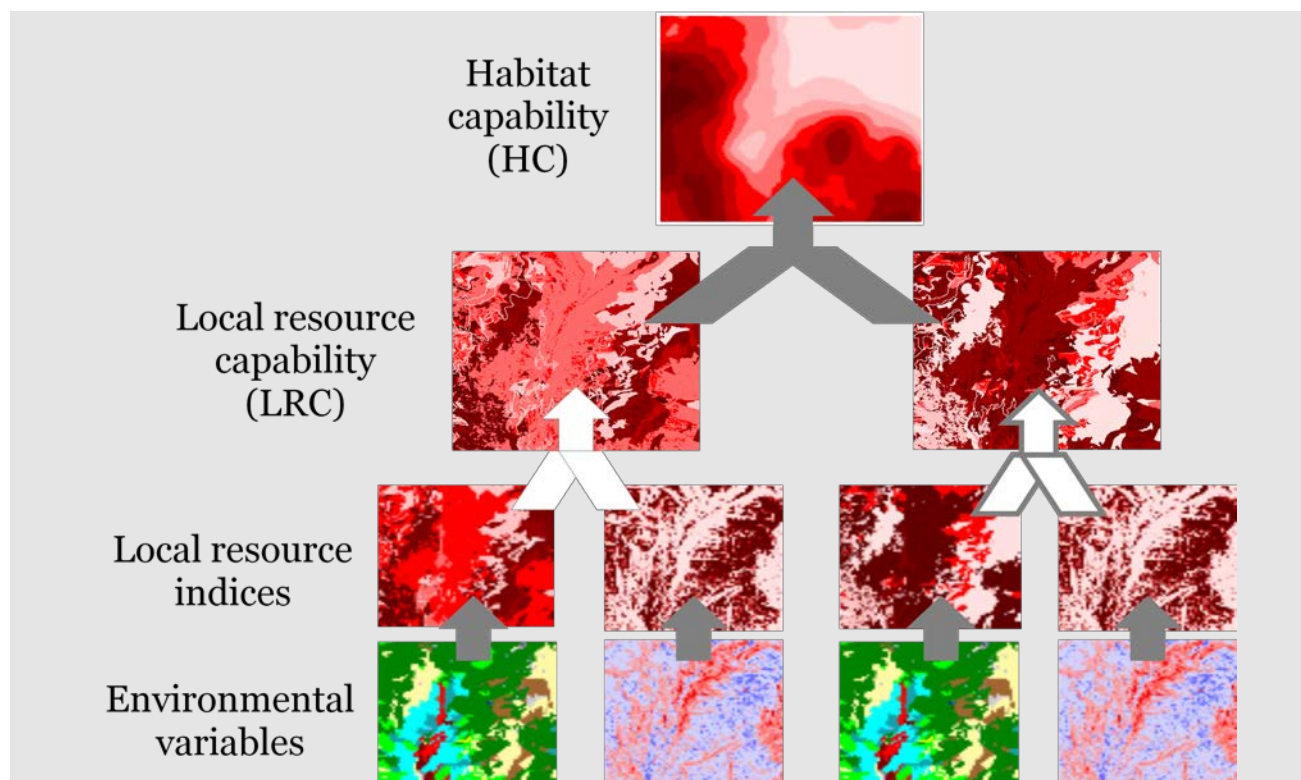


Figure 5. Schematic outline of the hierarchical structure of HABIT@ models.

Environmental variables represented as GIS layers are converted into local resource indices, which are combined to represent the availability of one or more local resources (e.g., cover, forage). The quantity, quality and accessibility of local resources within a potential home range centered on a focal cell determines the Habitat Capability (HC) of each focal cell. Note, the extent of suitable habitat within a neighborhood extending over multiple home ranges is incorporated directly into the calculation of HC to partially account for population-level considerations.

- 2) *Habitat Capability (HC)* — the capability of an area corresponding to an individual's homerange to support an individual, based on the quantity, quality and accessibility of local resources; habitat capability is a function of the amount and accessibility of high-quality local resources in the neighborhood of the focal cell; it is fundamentally an attribute of the local neighborhood and not the focal cell, even though the assessed value is returned to the focal cell. Note, an *HC* model can also include a coarse-scale consideration of habitat extent beyond the scale of a single homerange, for example based on the quantity, quality and accessibility of habitat within a much larger neighborhood of the focal cell corresponding to some multiple of the average homerange size. The coarse-scale extent of habitat often influences local population sizes and thus can influence the availability of mates at the local scale. If mate availability is viewed as a local resource (albeit a conceptual stretch), we can at least partially incorporate coarse-scale population level considerations into the derivation of *HC*. Thus, an isolated homerange of high habitat quality can receive a lower score than an area containing multiple homeranges of high quality.

Like an *HSI* model, HABIT@ returns a value between 0 (no habitat value) and 1 (prime habitat) for each cell at each level of the model. At the local resource level, the cell *LRC* value indicates the resource value at that cell (e.g., value as nesting habitat or forage) and there is a separate value for each designated local resource (which varies among species), although these grids are stored only as interim products. At the homerange level, the cell *HC* value indicates the relative value of a potential homerange centered on that cell with consideration of its broader neighborhood context. Note, *HC* values are continuous across the landscape, thus no assumptions are made as to where individual homeranges would actually be placed (**Fig. 6**).

Importantly, *HC* is not an estimate of occupancy or any specific population parameter (e.g., survival, productivity), since the latter depend on many other factors besides habitat (e.g., demographics). Rather, *HC* is an index of habitat capability that should vary with individual and population performance to the extent that habitat drives the latter. Thus, an index of say 0.8 at one location indicates that habitat conditions in the neighborhood of that location are estimated to be twice as good as another location that has an index of 0.4. In addition, *HC* is a species-specific index. *HC* absolute values are not appropriate for comparison across species due to differences in the underlying species' models; instead, *HC* values are best used within a species to compare over time or among landscape change scenarios.

In the context of the LCAD model, we generate a map of *HC* for each species in 2010, and an average across landscape change simulations in 2080. However, for parsimony sake the *HC* grids are not distributed as separate products.

3.3.1 Spatial data

HABIT@ models are based on GIS grids representing a variety of environmental variables such as cover type, stand age, biomass, slope, hydrological regime, roads and development. In the context of LCAD, these environmental variables include a modified version of the ecological systems map, a suite of ecological settings variables, and a suite of ecological integrity metrics, as follows.

3.3.1.1 Ecological systems

The primary land cover grid used in HABIT@ models is a modified version of the map developed by The Nature Conservancy in which they comprehensively mapped wildlife habitat in the eastern U.S (Anderson 2010) based on NatureServe's ecological systems ("NatureServe Explorer: An Online Encyclopedia of Life." 2011), which we refer to as DSLland. A detailed description of our DSLland map is provided elsewhere (McGarigal et al 2017), but with regards to the use of DSLland in HABIT@ models, there are three important modifications:

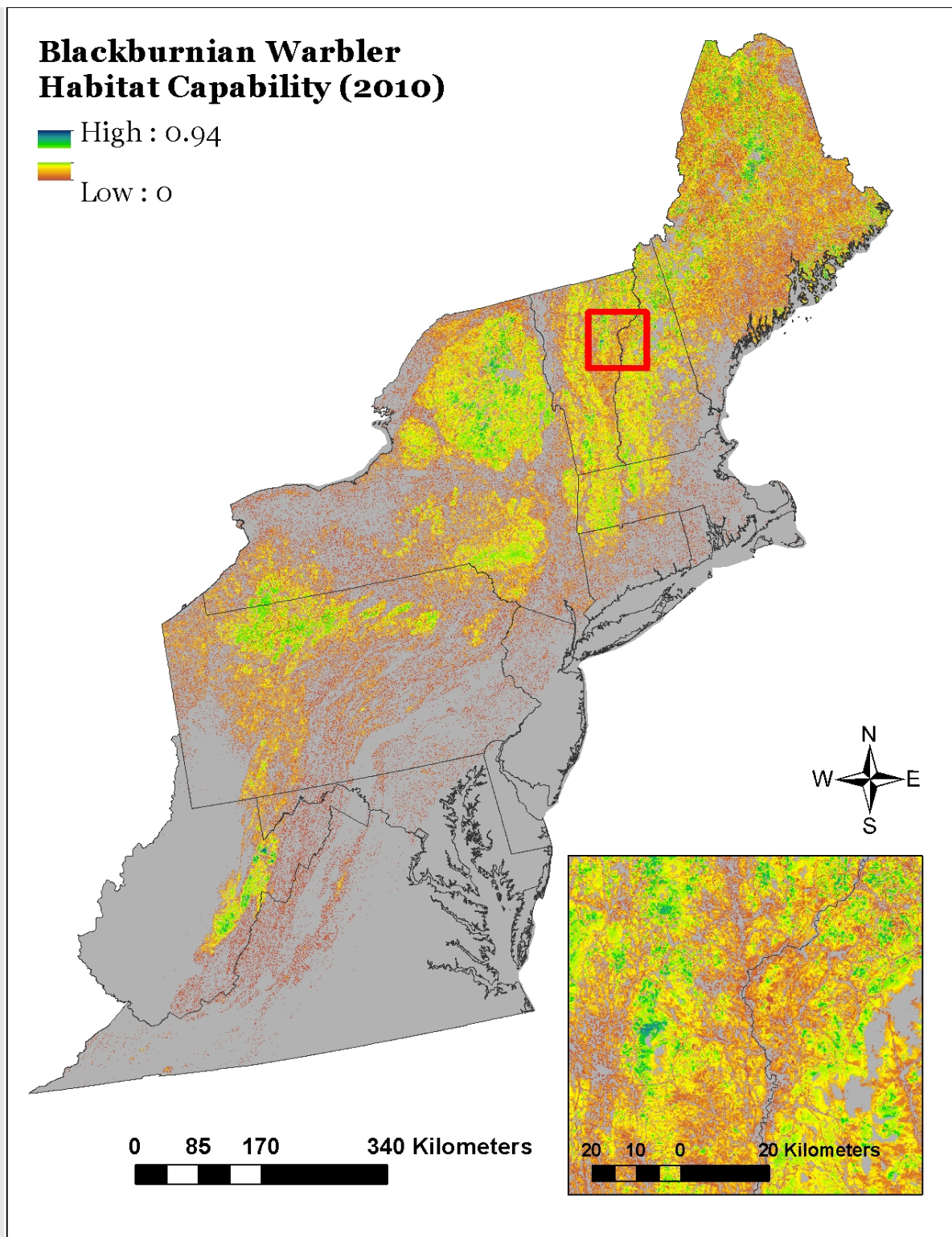


Figure 6. Example of a species' Habitat Capability (HC) index expressed as a continuous surface; shown here for the blackburnian warbler in the Northeast in 2010.

- 1) *Water bodies* — The original TNC map identifies all water as open water, and does not make distinctions between lentic and lotic. Furthermore, the TNC map does not map streams at all. These distinctions, as well as estimates of stream flow and gradient, are important components in assessing the capacity of the landscape to provide suitable habitat as measured by HABIT@, particularly for riparian-dependant species. Therefore, we replaced open water in the DSLland map with the National Wetlands Inventory (NWI) lentic and lotic classifications. Additionally, we used the NWI classifications for estuarine and marine tidal and subtidal classes. We also implemented a multi-step process to add National Hydrography Dataset (NHD) channel-forming streams with estimates of stream flow and gradient to the DSLland map.
- 2) *Development* — Although development was included in the TNC map, development intensity was not defined (i.e., the map does not differentiate between low, medium and high density development). Additionally, development in the TNC map was based on an out-dated data source, the 2001 National Land Cover Database (NLCD). We deemed both of these deficiencies as important in order to accurately assess the capacity of a landscape to support wildlife. Therefore, we burned into the DSLland map development classes from the 2011 NLCD grid (Xian et al. 2009) and maintained all development categories (low, medium and high).
- 3) *Roads and Traffic* — Roads were not explicitly part of the TNC map. They were represented in the TNC map only as developed cells (and thus completely confounded with development) as mapped by 2011 NLCD, were often omitted altogether due to their small footprint and linear configuration, and there was no contain information on road traffic rates. We deemed that accurate representation of roads, independent of development, as well as traffic rates, were crucial components to informative HABIT@ models. Therefore, in a complex, multi-step process, we removed roads from the developed class in the DSLland map and burned in new roads using the TIGER road shapefiles (“2010 TIGER/Line Shapefiles Main Page” 2011), and assigned traffic rates to each road segment.

3.3.1.2 Ecological settings variables

The ecological settings variables include a large suite of biophysical descriptors, including a suite of abiotic variables (e.g., soil texture, flow volume), biomass of forested systems, and anthropogenic variables (e.g., road traffic rate, terrestrial barriers). With respect to habitat modeling in particular, we selected total above ground live biomass (Mg/ha), as calculated using the equation from Jenkins et al. (2003), as the sole vegetation characteristic. We believed, *a priori*, that biomass is an important predictor of species habitat quality, providing information regarding seral stage, and therefore useful for building HABIT@ models.

Briefly, our current approach models biomass growth trajectories based on changes in biomass, as measured by Forest Inventory and Analysis (FIA; 1988-2011), given stand age and other site-level ecological setting covariates. These models are applied to the National Biomass and Carbon Dataset (NBCD) for the year 2000 (<http://www.whrc.org/mapping/nbcd/>) to grow biomass to the year 2010. Finally, recent

forest disturbances mapped by the University of Maryland are incorporated into the 2010 biomass estimates. The biomass growth models are then applied to the 2010 layer to in conjunction with stochastic disturbances to acquire 2080 biomass estimates for forested ecological systems. See technical document on disturbance and succession (McGarigal et al 2017) for a full explanation of this process.

3.3.1.3 Ecological integrity metrics

The ecological integrity metrics are associated with the coarse-filter assessment and include a suite of metrics computed for each cell representing the relative intactness (i.e., level of anthropogenic impairment) and resiliency to anthropogenic stress (see technical document on integrity, McGarigal et al 2017, for details). While these metrics are principally reserved for the coarse-filter assessment, they are available for use in HABIT@ models as well.

3.3.2 Resistant kernels

Resistant kernels play an important role in determining home range capability (and in the computation of the ecological integrity metrics) and are a relatively novel method, so they warrant special attention here. The resistant-kernel estimator, introduced by Compton et al. (2007), is a hybrid between two existing approaches: the standard kernel estimator and least-cost paths based on resistant surfaces.

The *standard kernel estimator*, given two-dimensional data (e.g., x,y points), produces a three-dimensional surface representing an estimate of the underlying probability distribution by summing across bivariate curves centered on each sample point. The standard kernel estimator begins by placing a standard kernel over each sample point. Here, we can think of the standard kernel as estimating the ecological neighborhood of a focal cell (or sample point). The sum of all the kernels is a surface that represents the density of sample points in the ecological neighborhood of any location. Standard kernels can be used to estimate density for point features (e.g., vernal ponds), linear features (e.g., roads), patches (e.g., land cover types), and even continuous surfaces (e.g., canopy cover). The key is that the standard kernel allows one to estimate the density of some feature of interest that incorporates some ecological knowledge about the size and shape of the ecological neighborhood.

Resistant surfaces (also referred to as cost surfaces) are being increasingly used in landscape ecology, replacing the binary habitat/nonhabitat classifications of island biogeography and classic metapopulation models with a more nuanced approach that represents variation in habitat quality (e.g., Ricketts 2001). In a patch mosaic, for example, a resistance value (or cost) is assigned to each patch type, typically representing a divisor of the expected rate of ecological flow (e.g., dispersing or migrating animals) through a patch type. For example, a forest-dependent organism might have a high rate of flow (and thus low resistance) through forest, but a low rate of flow (and thus high resistance) through high-density development. In this case, the cost assigned to each patch type in the resistant surface may represent the willingness of the organism to cross the patch type, the physiological cost of moving through the patch type, the reduction in survival for the organism moving through the patch type, or an integration of all these factors. Empirical data on costs are often lacking, but can be derived from a variety of data sources, including

detection, movement (e.g., capture/recapture, telemetry) and/or genetic data for the organism (or process) under consideration. Traditional least-cost path analysis finds the shortest functional distance between two points based on the resistant surface. The cost distance (or functional distance) between two points along any particular pathway is equal to the cumulative cost of moving through the associated cells. Least-cost path analysis finds the path with the least total cost. This least-cost path approach can be extended to a multidirectional approach that measures the functional distance (or least-cost path distance) from a focal cell to every other cell in the landscape, or from every cell (treated as a focal cell) to every other cell. In this sense, the multi-directional approach (from all cells to all cells) represents the most synoptic approach available for measuring functional connectivity.

In the *resistant kernel estimator*, the complement of least-cost path distance (a.k.a. functional proximity) to each cell from the focal cell is multiplied by a weight reflecting the shape and width of the standard kernel. Consequently, given the typical shape of a standard kernel (e.g., Gaussian), the least-cost path distance asymptotically approaches infinity after roughly three standard deviations from the focal cell. The end result is a resistant kernel that depicts the functional ecological neighborhood of the focal cell (**Fig. 7**). In essence, the standard kernel is an estimate of the fundamental ecological neighborhood and is appropriate when resistant to movement is irrelevant (e.g., highly vagile species), while the resistant kernel is an estimate of the realized ecological neighborhood when resistance to movement is relevant. The resistant kernel can also be thought of as representing a process of spread (e.g., dispersal) outward from the focal cell, that combines the cost of moving through a heterogeneous resistant neighborhood with the inherent typically nonlinear cost of moving any distance away from the focal cell.

Resistant kernels have a wide variety of uses in ecological modeling. In the context of HABIT@, we use a resistant kernel approach for computing home range capability. Specifically, resistance values are assigned to land cover types such that home range kernels preferentially expand into higher quality land cover types that are assigned low resistance values. Barriers to home ranges (e.g., developed areas) are assigned high resistance values. Resistance values are species-specific and represent an organism's willingness to incorporate different cover types into the home range. The least-cost path analysis yields the functional proximity of surrounding cells to each focal cell. Scaling the proximities by a density function (Gaussian in this case) results in a resistant kernel for a focal cell. The resulting resistant kernel is a three-dimensional surface with a volume equal to a value less than or equal to one (a resistant kernel in a homogeneous non-resistant neighborhood has a value of one) representing the probability of use surrounding the focal cell; in other words, it represents a potential home range utilization kernel centered on the focal cell (**Fig. 7**). This home range kernel is used to weight the local resources available at each cell within the potential home range centered on the focal cell to derive an index of habitat capability for the focal cell. The use of the resistant kernel means that the habitat capability index represents not only the quantity and quality of local resources surrounding the focal cell, but also the accessibility of those resources. In HABIT@ models, we build a resistant kernel (i.e., home range kernel) and assess habitat capability for each cell that confers any breeding habitat value.

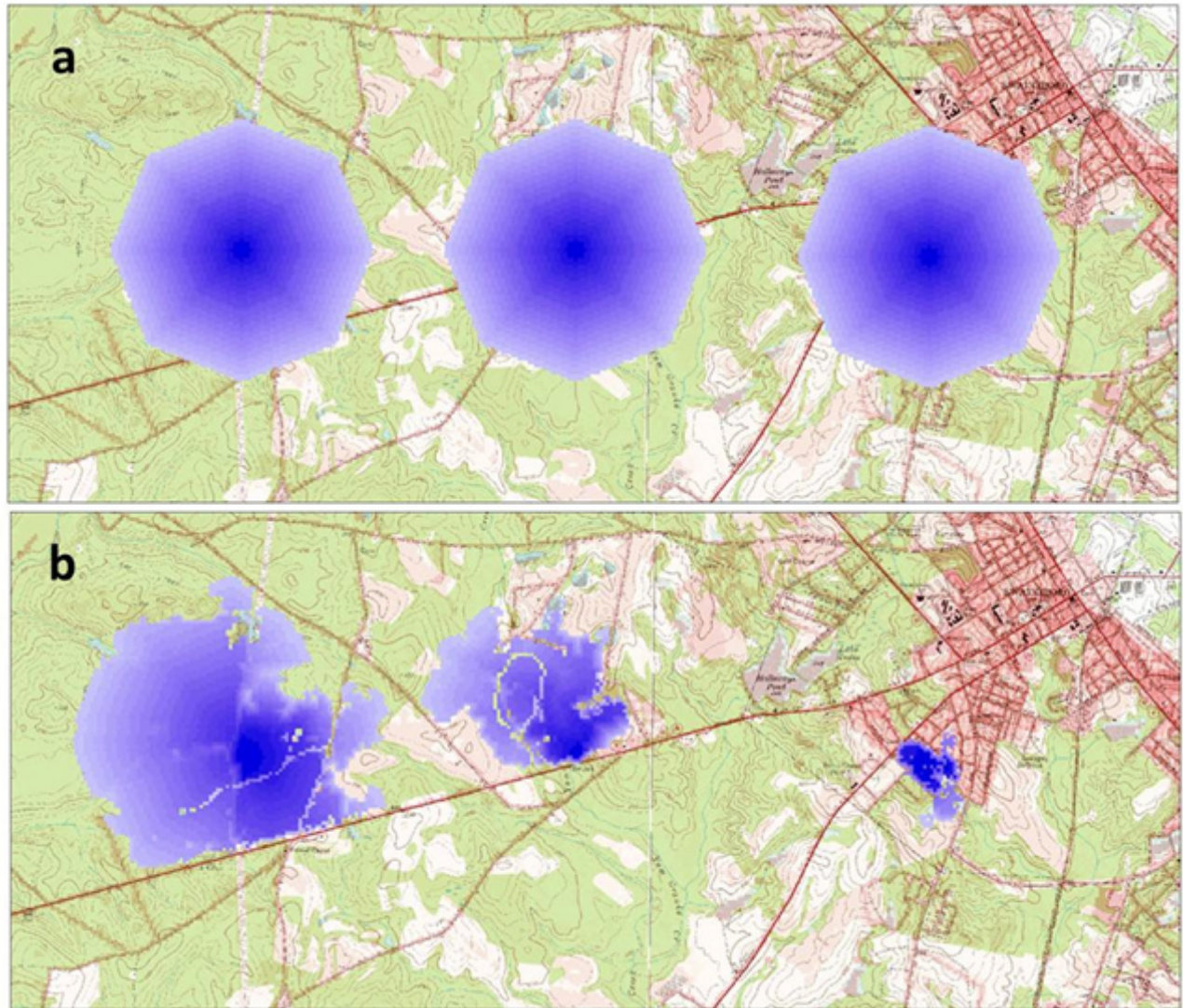


Figure 7. Hypothetical resistant kernel used to depict a potential home range utilization distribution centered on a focal cell. The resistant kernel spreads outward from the focal cell iteratively depleting a "bank account" based on local resistance (e.g., as conferred by various land cover types, in this example-developed) until the bank account is completely deleted. The bank account is scaled to the desired bandwidth of the kernel. The kernel value represents the least-cost path distance to each cell from the focal cell. The complement of the least-cost path distance (a.k.a. proximity distance) is multiplied by a weight reflecting the shape and width of the standard kernel (usually based on a Gaussian function) and the result is a resistant kernel that reflects local resistance and ecological distance relationship expressed via the size and shape of the standard kernel. Shown here is the difference between a standard kernel (a) and a resistant kernel (b) for three focal cells in a heterogeneous landscape.

3.4 Prevalence Model

The next step is to develop a *Prevalence* model for each selected representative species. Briefly, a prevalence model estimates how prevalent a species is (i.e., its probability of occurrence) across its range based solely on its current spatial distribution without explicit regard to climate suitability, habitat capability or other factors. In the context of the LCAD model, we use the species' prevalence model to account for unmeasured biogeographic factors influencing the species' distribution that is not necessarily captured by the climate niche or habitat model. Importantly, prevalence is estimated for the current timestep (2010) only and is thus treated as static for the 70-year simulation, in the same manner as the distribution of ecological systems. A general description of the prevalence modeling approach follows and is illustrated using results for wood thrush. The process described next is the same for all species, even though some of the details, such as sources of empirical data, may vary slightly among species.

First, we compile an independent data set on species occurrence within the assessment area. Our goal is to consistently use Breeding Bird Survey (BBS) data (as described above) to derive prevalence models for bird species. However, in the event that BBS data does not provide adequate samples, for example with non-avian species or avian species with very restricted distributions, we augment this dataset with other existing data as needed and deemed appropriate (e.g., Mountain Birdwatch data).

Second, we subdivide the landscape (i.e., Northeast region) into 20 km square cells and compute a Gaussian distance-weighted interpolation of proportional presence using a 5 km bandwidth (standard deviation) for the Gaussian kernel. Briefly, this interpolation process produces an estimate for each 20 km cell by computing the average proportional presence among all neighboring observations (e.g., BBS segments) after weighting each observation by its Gaussian distance (from the centroid of the observation) to the focal cell.

The output of the prevalence model is a continuous surface grid depicting prevalence for each species in 2010; values range between 0 (no occurrences within say 20 km) and 1 (100% present at nearby observations) (**Fig. 8**). Note, the prevalence grid is not distributed as a separate product.

3.5 Landscape Capability Model

The next step is to integrate the climate niche, habitat capability and prevalence into a single *Landscape Capability (LC)* model for each selected representative species. Briefly, a *LC* model represents an index of the species' relative probability of occurrence based on climate, habitat and other biographic factors (as represented by prevalence); it is our best estimate of the species' distribution based on measurable factors (**Fig. 9**). Note, we acknowledge that there may be other factors influencing species' distributions, such as interspecific interactions, human persecution and disease, but these are outside the scope of modeling with extant data at the regional scale. In the context of the LCAD model, we use the species' *LC* to predict the current and future distribution of the species under alternative landscape change scenarios. Importantly, these predictions are used to determine where the species might occur based on climate suitability, habitat capability

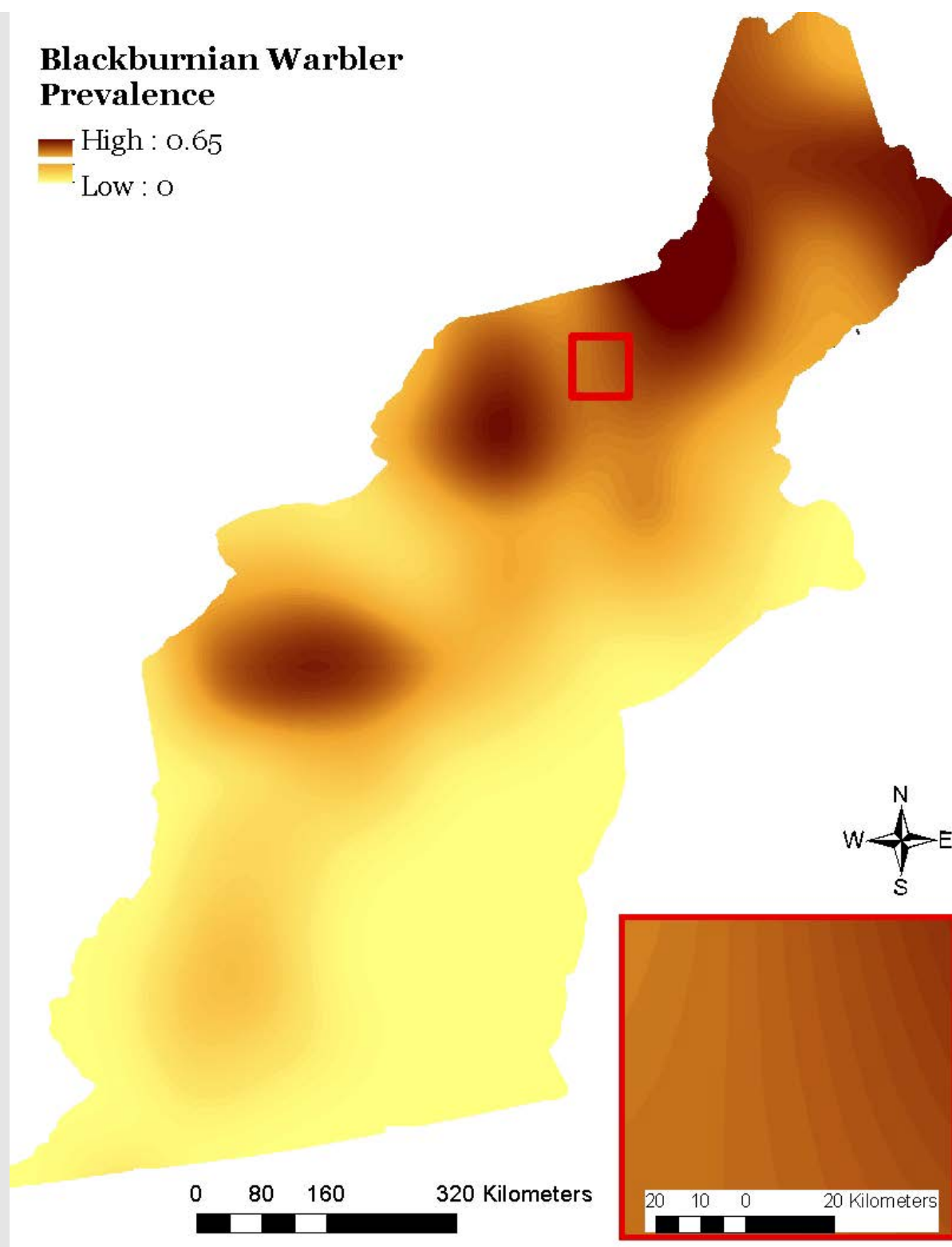


Figure 8. Example of a species' Prevalence index expressed as a continuous surface; shown here for the blackburnian warbler in the Northeast in 2010.

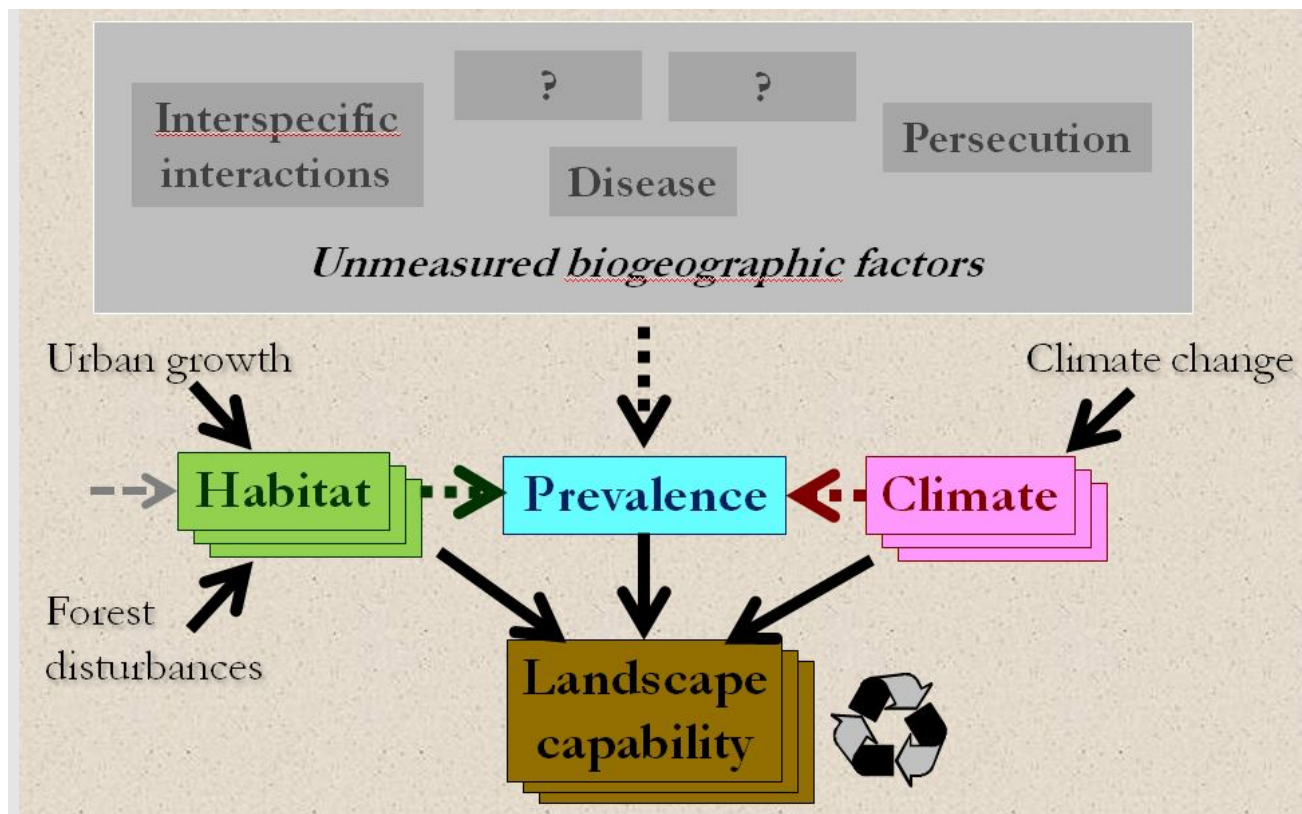


Figure 9. Schematic outline of the landscape capability modeling framework, in which we separately model the species' climate niche, habitat capability and prevalence, while recognizing that there are potentially other factors influencing the species' distribution, and then integrate these factors into single index.

and prevalence, not where they will actually occur, since the latter depends on whether they can and choose to get there from currently occupied sites and the influence of other unmeasured factors. A general description of the *LC* modeling approach follows and is illustrated using results for wood thrush.

To compute *LC*, we simply multiply the *CN*, *HC* and Prevalence grids together that essentially treats each of three factors as potentially limiting factors. In other words, by multiplying the grids together, we ensure that if a cell has a low value for any of the grids, the resulting product will too be low. This safeguards, for example, against locations with suitable climate and habitat but no prevalence from receiving a high relative probability of occurrence in the *LC* model. Similarly, it ensures that a location with suitable climate and within the species range but containing no habitat will not receive a high relative probability of occurrence in the *LC* model.

In some cases, where BBS and other supplemental datasets such as MBW do not adequately sample the full extent of the species' range (e.g., blackpoll warbler), the Prevalence estimate can exclude portions of the range, thus erroneously reducing the surface to near 0 at these locations. In such cases, we exclude prevalence from the *LC* index and instead simply multiply the *CN* and *HC* surfaces together to produce the *LC* index.

Based on the approach described above, we generate the following spatial index:

- *Current Landscape Capability (speciesLC)* — map depicting each species' relative probability of occurrence scaled 0-1 based on the integrated climate-habitat-prevalence index in 2010 (**Fig. 10**).

3.5.1 LC model evaluation

Because the *LC* model described above is not a statistical model per se, we cannot evaluate the model internally using conventional measures. Instead, we evaluate the performance of the model against independent data using a variety of statistical measures, which ultimately provides us with a true model validation, as described below.

First, for species with suitable empirical data on point-level occurrence, we compile an independent data set on occurrence within the assessment area. Note, for the purpose of *LC* modeling, coarse-scale data such as the BBS (8 km route segments) and Breeding Bird Atlas (5 km squares) is not sufficient, since our purpose is to associate occurrence at the 30 m cell level. Our goal is to consistently use eBird data (eBird Reference Dataset Version 3.0 from the Avian Knowledge Network; <http://www.avianknowledge.net/content/features/archive/ebird-reference-dataset-3-0-released>) to evaluate *LC* models for bird species. eBird is a dataset maintained by the Cornell Lab of Ornithology and the National Audubon Society of count data for bird species submitted by novice and expert observers (Munson et al. 2011). Note, as mentioned above, for several bird species BBS is not adequate to develop CN models; for these species we use eBird data to develop the CN model. Unfortunately this precludes eBird as a source of data for model evaluation for these species. In these cases and for non-bird species, we attempt to locate other independent, regional data to evaluate the *LC* models.

For suitable species, we restrict eBird data in the following ways:

- 1) We use data from 2005 – 2010, as this time period best matches our land cover data.
- 2) We limit surveys to those that occurred only during daylight hours between 05:00 and 20:00 (Fink et al. 2010).
- 3) We limit surveys to dates corresponding to the seasonality of the species' habitat capability model. For example, for a breeding season model, we limit the surveys to those dates when detections can safely be considered those of breeding, not migrating individuals; e.g., Julian dates 145-222 (May 25-August 10) for the wood thrush model.
- 4) We limit surveys to "stationary counts" (i.e., those conducted from a single location) to ensure that observations can most reliably be attributed to a single location within the scale of the climate and habitat data).
- 5) To reduce the influence of varying survey time on detections, we limit surveys to those <3 hours in duration (Fink et al. 2010).
- 6) We prioritize the use of "complete checklist" survey types. For "complete checklists" participants are asked to submit all birds detected during the survey. Therefore it is assumed that species not listed on the checklist are absent (Fink et al. 2010). Complete checklists provide a sample of "presences" and "absences", but typically the

**Blackburnian Warbler
Landscape Capability(2010)**

High : 0.45

Low : 0

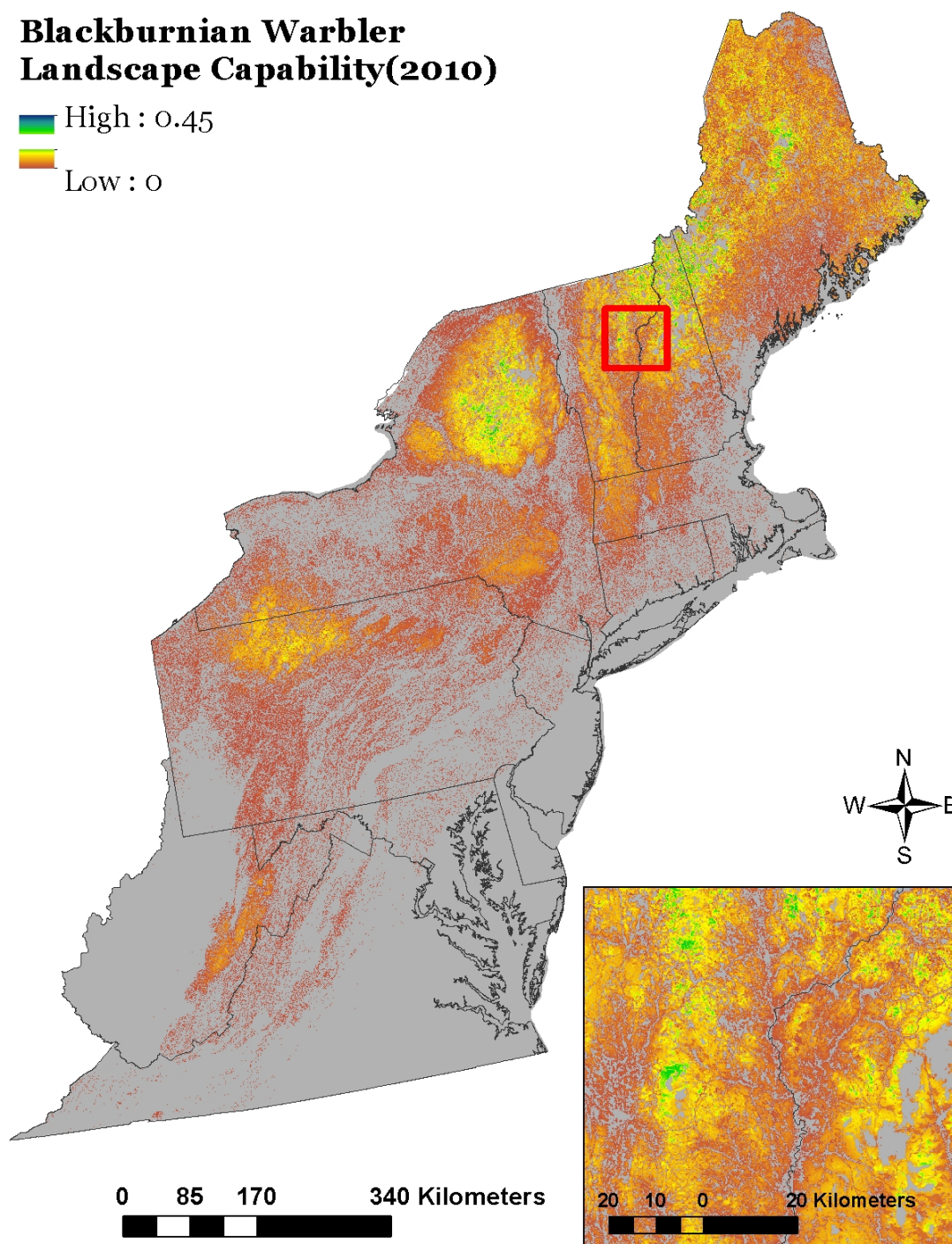


Figure 10. Example of a species' current Landscape Capability (LC) index, expressed as a continuous relative probability of occurrence surface; shown here for the blackburnian warbler in the Northeast in 2010.

"absences" far outnumber the "presences" and often the total number of "presences" is deemed inadequate to build reliable models. Consequently, we supplement the

- 7) "presences" with "present" observations from the "basic" surveys (which are deemed present-only surveys and thus absences cannot be implied) in an attempt to produce an equal number of confirmed "presences" and "absences".
- 8) To reduce the likelihood of selecting late or early migrants or otherwise abnormal observations, when selecting "present" locations from the basic surveys (to supplement those from the checklists), we constrain the selection to observations within the climate niche envelope (CNE).
- 9) To reduce the influence of residual spatial autocorrelation, when selecting observations from the checklist and basic surveys, we constrain the selection process so that only one location is selected within a 250 m radius.
- 10) To reduce the influence of geographic bias in the sampling, for each eBird sample location that meets the criteria above, we weight each observation by the sampling intensity within a 25 km window. Specifically, to avoid the bias caused by unevenly distributed survey effort across the region, we impose a 25 km grid on the assessment area and count the number of survey points in each grid cell. Each survey point is weighted by the reciprocal of the number of sample points in the corresponding cell, effectively resulting in equal weight assigned to each grid cell (since the sum of weights in each cell equals one). Note, this is a computationally efficient substitute for computing the reciprocal of the kernel density of points around each survey location. Note, during preliminary analyses we imposed grids of 25, 50, 75, and 100 km and found the results of the statistical analysis relatively insensitive to the grid resolution, so we retained the finest grid resolution (25 km) to preserve the largest effective sample size.

When there is a large discrepancy in the number of presences and absences a random sample of the larger group is taken, such that the sample of presences is equal to the sample of absences.

Second, to validate the predictive ability of the *LC* index, we use simple logistic regression to predict the probability of occurrence as a function of the *LC* index, with the following modeling considerations: 1) we treat each observation as an independent observation and make no formal attempt to account for residual spatial autocorrelation (other than as noted above), as it is not deemed critical to the point estimates of model parameters (and thus model predictions); 2) we assume a binomial error distribution for the binary response; 3) we include survey date and time as detection covariates and include effort hours as an offset in the model (note, unfortunately the survey design does not allow for a more formal treatment of detectability in the model); 4) we weight each observation by the reciprocal of the local sampling intensity (as described above); and 5) we take the maximum value of *LC* within 250 m of the survey point as the value for the predictor to account for two factors: a) that birds, for example, can often be detected at considerable distances from survey points, and b) georeferencing errors in the survey locations (which we know exist in these datasets). Note, the regression model predicts the probability of occurrence for each observation; a cutpoint or threshold value on the response scale (i.e., probability scale)

must be selected in order to classify observations as "present" or "absent". In our case, we select the threshold that maximizes the chance-corrected correct classification rate based on the Kappa statistic. This approach chooses a threshold that balances the errors of omission and commission.

Third, to assess *LC* model performance, we compute several summary statistics based on the logistic regression above, with an emphasis on the model's predictive ability:

- 1) *Deviance explained* — percent of the deviance explained by the *LC* model. Deviance explained is a measure of the variance in the response explained by the generalized linear model and is analogous to R^2 in linear regression. High percent deviance explained indicates a good model -- one that is able to explain the variation in the response well. Note, deviance explained is not a measure of the model's predictive ability, but rather its explanatory power.
- 2) *Area under the curve* (AUC) — area under the Receiver Operator Characteristic (ROC) curve. The ROC curve describes the compromises that are made between the model sensitivity (proportion of observed presences correctly predicted to be present) and false positive (observed absences wrongly predicted to be present) fractions as the decision threshold is varied; a species is predicted to be present or absent at a site based on whether the predicted probability for the sample is higher or lower than a specified threshold probability value. Furthermore, the sensitivity and false positive values are independent of the prevalence of the positive group (i.e., the proportion of total observations in the present group) because they are expressed as a proportion of all sites with a given observed state (Swets, 1988). ROC plots provide a threshold-independent method of evaluating the performance of presence/absence models. A good model will achieve a high true positive rate while the false positive rate is still relatively small; thus the ROC plot will rise steeply at the origin, and then level off at a value near the maximum of 1. The ROC plot for a poor model (whose predictive ability is the equivalent of random assignment) will lie near the diagonal, where the true positive rate equals the false positive rate for all thresholds. ROC analysis is therefore independent of both species prevalence and decision threshold effects, which gives rise to the AUC index.

AUC ranges from 0.5 for models with no discrimination ability, to 1 for models with perfect discrimination. This index can also be interpreted in terms of the true positive and false positive values used to create the curve. A general rule of thumb for interpreting the index is as follows. Areas between 0.5 and 0.7 indicate poor discrimination capacity because the sensitivity rate is not much more than the false positive rate. Values between 0.7 and 0.9 indicate a reasonable discrimination ability appropriate for many uses, and rates higher than 0.9 indicate very good discrimination because the sensitivity rate is high relative to the false positive rate (Swets, 1988). Hanley and McNeil (1982) have shown that the ROC index can also be interpreted as the probability that a model will correctly distinguish between two observations, one positive and the other negative. In other words, if a positive observation and a negative observation are selected at random the index is an estimate of the probability that the model will predict a higher likelihood of occurrence for the positive observation than for the negative observation.

- 3) *Kappa* — chance-corrected correct classification rate. The discrimination performance of a binary logistic regression can be assessed by examining the agreement between predictions and actual observations, using a 2x2 classification table (also known as a confusion matrix); a species is predicted to be present or absent at a site based on whether the predicted probability for the sample is higher or lower than a specified threshold probability value. For a specified threshold value, *Kappa* measures the proportional improvement in correct classification over that expected by chance. Note, unlike AUC, *Kappa* is a threshold-dependent measure of model discrimination. Here, we select the threshold that maximizes *Kappa*, and judge model performance by the magnitude of this *Kappa*.

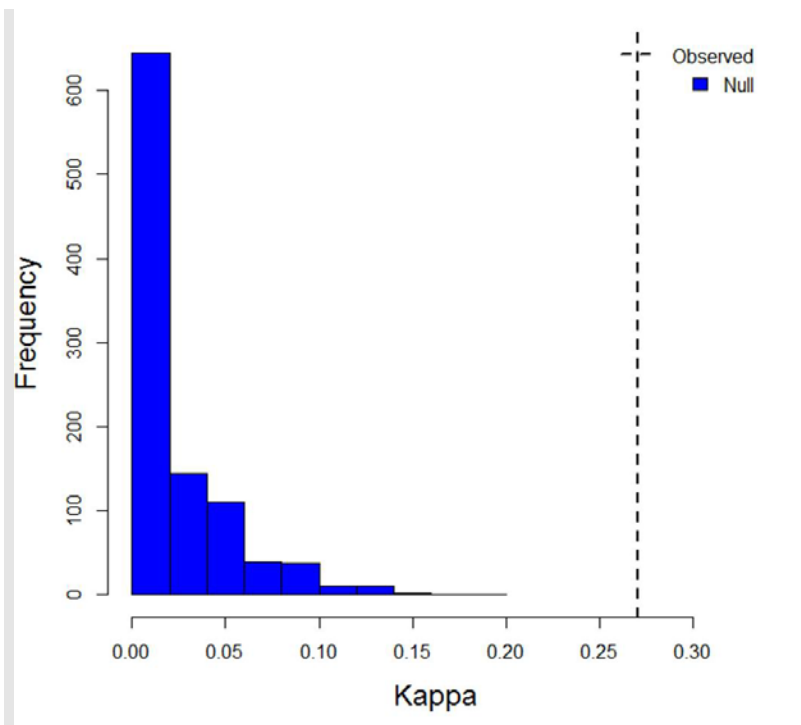


Figure 11. Monte Carlo randomization test (1,000 replications) Of the observed *Kappa* (chance-correct correct classification rate) for the logistic regression model predicting species' occurrence from an integrate index of climate niche, habitat capability and prevalence; shown here for blackburnian warbler in the Northeast. The null distribution of *Kappa* is given by the frequency distribution, while the observed *Kappa* (0.27) is shown as a vertical dashed line. The *p*-value (<0.001, in this case) is computed as the proportion of the null distribution greater than or equal to the observed *Kappa*.

We also conduct a Monte Carlo randomization test of the observed maximum *Kappa*. Specifically, we compute the observed *Kappa* and compare it to the distribution of *Kappa* under the null distribution (**Fig. 11**). The null distribution represents the distribution of *Kappa* expected if the species' occurrences are distributed randomly with respect to maximum *LC* values with 250 m of the survey point (i.e., our *LC* index has no real relationship to the species' distribution). We generate the null distribution by randomly shuffling the "present" and "absent" locations, recomputing *Kappa*, and repeating the process 1,000 times. We expect the original model to have a greater *Kappa* than the bulk of the null distribution if the species' distribution is nonrandom with respect to maximum *LC*. Specifically, we compute the upper (one-sided) *p*-value as the proportion of the null distribution greater than or equal to the observed *Kappa*.

We interpret a significant difference (i.e., $p\text{-value} \leq 0.05$) as a confirmation of the model's discriminatory power.

- 4) *Omission and Commission error rates* — at the threshold value that maximizes Kappa, the resulting confusion matrix gives the number of observed presences wrongly predicted to be absent (omission errors) and the number of observed absences wrongly predicted to be present (commission errors). Note, like Kappa, the omission and commission error rates are a threshold-dependent measure of model discrimination. Here, we report the omission and commission error rates for the threshold that maximizes Kappa and judge model performance by the magnitude of these error rates.

3.6 Landscape Change Indices

The final step is to compute a variety of non-spatial and spatial indices to reflect the species' potential response to climate change, habitat change or both.

First, we compute a complementary set of non-spatial indices for each species based on the proportional change in *LC* due to climate change, habitat change, or both within the specified geographic extent. These non-spatial indices are primarily useful for establishing conservation objectives or targets for species in conservation design or for comparison among landscape change scenarios. The response indices are computed for each landscape change simulation based on the 2080 predictions and summarized as the mean across simulations, and the "current" below refers to 2010 (**Table 3**).

- 1) *Current LC* — sum of current range-rescaled *LC* values multiplied by 0.09 to convert it to maximum *LC*-equivalent hectares; it ranges from 0 (no habitat) to the study area extent in hectares (when the entire study area is all optimal habitat). Note, this is a baseline metric for computation of the other indices.
- 2) *Future climate response* — sum of future raw-scaled *LC* values calculated with current habitat and predicted future climate in 2030 and 2080 (averaged across RCP4.5 and 8.5 scenarios), minus the sum of current (2010) raw-scaled *LC* values, divided by the sum of current (2010) raw-scaled *LC* values, multiplied by 100 to convert to a percentage; it ranges from -100 (complete loss of *LC* due to climate change) to unbounded positive values reflecting an expanding *CNE* within the project area. This index represents the percentage change in *LC* from 2010 to 2080 due solely to climate change.
- 3) *Future habitat response* — sum of future raw-scaled *LC* values calculated with current climate and predicted future habitat in 2030 and 2080 (averaged across urban growth simulations), minus the sum of current (2010) raw-scaled *LC* values, divided by the sum of current (2010) raw-scaled *LC* values, multiplied by 100 to convert to a percentage; it ranges from -100 (complete loss of *LC* due to habitat change) to unbounded positive values reflecting improving habitat conditions within the project area. This index represents the percentage change in *LC* from 2010 to 2080 due solely to habitat change.

- 4) *Future climate and habitat response* — sum of future raw-scaled *LC* values in 2030 and 2080 (averaged across urban growth and climate change simulations) minus the sum of current (2010) raw-scaled *LC* values, divided by the sum of current (2010) raw-scaled *LC* values, multiplied by 100 to convert to a percentage; it ranges from -100 (complete loss of *LC* due to landscape changes) to unbounded positive values reflecting improving *LC* within the project area. This index represents the percentage change in *LC* from 2010 to 2080 due to both climate and habitat changes.

Second, we derive a variety of spatial indices representing the species' potential response to climate change, habitat change or both based on changes in *LC* under different assumptions or for different purposes. These spatial indices are useful either for prioritizing locations for conservation action for each species in the context of landscape conservation design or for visualizing the potential changes in the distribution of a species due to climate change, habitat change or the combination of the two. These spatial indices are computed for each landscape change simulation based on the 2080 predictions only (for parsimony sake) and summarized as the mean across simulations.

- 1) *Future Landscape Capability (speciesLC2080)* -- map depicting each species' index of occurrence based on an average of the predicted *LC* surface across landscape change simulations in 2080 (**Fig. 12**). Specifically, we derive the *LC* index in 2080 under each landscape change simulation (i.e., urban growth and climate change simulation) and average the predictions to depict the future *LC* as a continuous surface that ranges from 0 (non habitat) to 1 (optimal conditions).
- 2) *Climate response (speciesCR2080)* -- map depicting the future *LC* calculated with current habitat and predicted future climate in 2080 (averaged across RCP4.5 and RCP8.5 scenarios) within the project area (**Fig. 13**). This index emphasizes places with high current habitat and climate capability that maintain or increase in climate suitability over time without regard to future changes in habitat capability.
- 3) *Habitat response (speciesHR2080)* -- map depicting the mean future *LC* calculated with current climate and predicted future habitat in 2080 (averaged across urban growth simulations) within the project area (**Fig. 14**). This index emphasizes places with high current habitat and climate capability that maintain or increase in habitat capability over time without regard to future changes in climate suitability.

DSL Project Component: Modeling focal species

Table 3. Example species' landscape change indices in the Northeast Region in 2010. Current (2010) Landscape Capability (LC) is equal to the LC index summed across cells in 2010, expressed in hectares, and it ranges from 0 (no habitat) to the extent of the project area (i.e. when the entire project area is optimal habitat). Landscape change indices represent the percentage change in LC from 2010 to 2030 or 2080 due to climate change, habitat change or both (see text for details); 0 reflects no change from the current.

Species	Current (2010) Landscape Capability (LC)	Landscape Change Indices (%)					
		Climate Response		Habitat Response		Climate & Habitat Response	
		2030	2080	2030	2080	2030	2080
American woodcock	72,916,983	-1.1	-6.1	tbd	tbd	tbd	tbd
Blackburnian warbler	65,756,067	-18.8	-68.0	tbd	tbd	tbd	tbd
Blackpoll warbler	1,541,033	-20.0	-65.1	tbd	tbd	tbd	tbd
Eastern meadowlark	4,745,130	18.0	46.3	tbd	tbd	tbd	tbd
Louisiana waterthrush	441,392	31.1	123.8	tbd	tbd	tbd	tbd
Marsh wren	1,598,923	17.0	62.0	tbd	tbd	tbd	tbd
Moose	289,266,365	-3.2	-4.4	tbd	tbd	tbd	tbd
Northern waterthrush	27,579,842	-30.2	-70.3	tbd	tbd	tbd	tbd
Prairie warbler	66,780	15.5	42.2	tbd	tbd	tbd	tbd
Ruffed grouse	166,875,203	7.8	-46.0	tbd	tbd	tbd	tbd
Wood duck	454,954	17.4	59.2	tbd	tbd	tbd	tbd

Figure under development

Figure 12. Example of a species' future Landscape Capability (LC) index, expressed as a continuous relative probability of occurrence surface averaged across landscape change simulations; shown here for the blackburnian warbler in the Northeast in 2080.

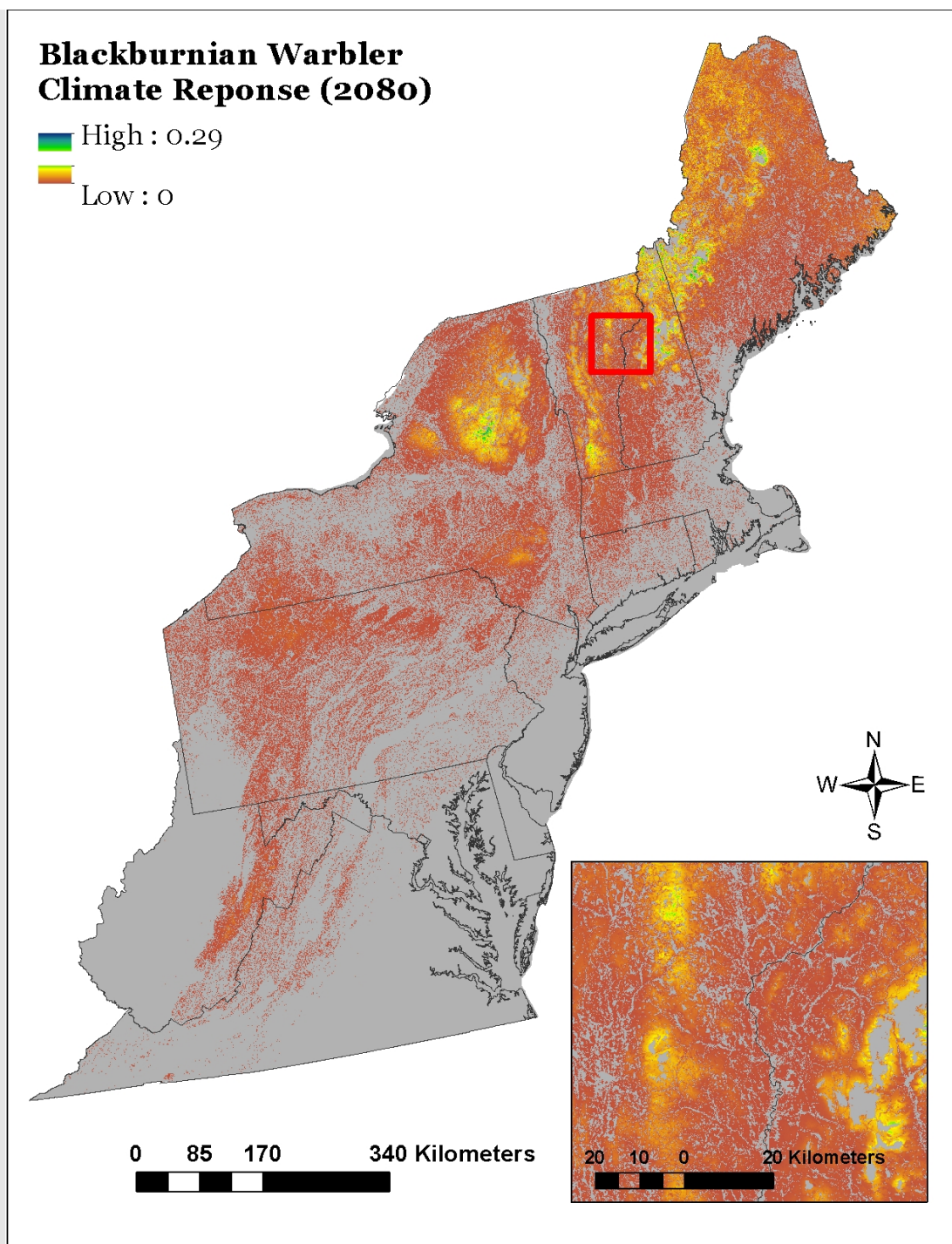


Figure 13. Example of a species' Climate Response index, defined as the future Landscape Capability index calculated with current habitat and predicted future climate in 2080 (averaged across RCP4.5 and RCP8.5 scenarios) within the project area. High values represent places with high current LC that maintain climate suitability over time without regard to future changes in habitat capability. Shown here for the blackburnian warbler in the Northeast.

Figure under development

Figure 14. Example of a species' Habitat Response index, defined as the future Landscape Capability index calculated with current climate and predicted future habitat in 2080 (averaged across urban growth simulations) within the project area. High values represent places with high current LC that maintain habitat capability over time without regard to future changes in climate suitability. Shown here for the Blackburnian warbler in the Connecticut River watershed.

4 Alternatives Considered and Rejected

We considered the following alternative and options:

First, we sought to develop multi-level occupancy (observation-process) models that account for detectability, but determined that the extant datasets available to us did not have a study design that lent itself to modeling detectability at the spatial resolution needed for LCAD (i.e., 30 m). In the statistical evaluation of *LC* we incorporated detection covariates (e.g., date and time) as fixed effects in the logistic regression model as the best available option to account for factors influencing detectability.

Second, we considered using a more conventional approach to model species-habitat relationships involving regressing species presence/absence on the raw habitat covariates. The pros of using this approach is that it is entirely empirically based (i.e., there is no expert opinion involved except in the selection of covariates to include in the model) and the results are straightforward and intuitive, and it represents the norm in the scientific literature. The cons of this approach is that it does not easily accommodate complex, multi-scale habitat relationships that account for habitat quantity, quality and accessibility. Instead, we opted to use HABIT@ models for following reasons. First, HABIT@ models can be constructed for any species regardless of data availability, because they are expert derived, and thus provide a consistent platform for modeling all representative species. Second, and more importantly, HABIT@ models allow for much richer habitat relationships, particularly multi-level and multi-scale relationships that account for the habitat quantity, quality and accessibility at multiple levels of organization (e.g., local resource, homerange and population), that better correspond to our understanding of species' habitat selection than conventional models. Third, the output of a HABIT@ model is merely a single integrated habitat covariate that is derived by an expert-driven process, but this is no different than deriving habitat covariates from raw data sources that is common in conventional habitat modeling; the difference is in the complexity of our derivation of the habitat covariate. Lastly, because we can assess and validate the final model using extant data on species' presence/absence, we can verify that the model performs well in light of the empirical data. Thus, while the parameters of the HABIT@ models are not all formally optimized with respect to the empirical data, the model is empirically assessed, which allows us to confirm that the model is a "good" model even if we cannot claim that it is the "optimal" model. Theoretically, it is conceivable to fully optimize a HABIT@ model based on empirical data, but the computational demands for such an undertaken are too great to make this practical with current computing resources.

Third, one area that involved substantial consideration of alternatives was how to incorporate climate change into species distribution modeling. Our original concept was to assess the influence of climate change on habitat capability by modeling shifts in the distribution of ecological systems, which then indirectly through the HABIT@ models would affect habitat capability. Ultimately, we concluded that it was not realistic to model changes in the distribution of ecological systems (see Ecological Systems documentation for more details). Given the static distribution of ecological systems, we considered integrating climate into the HABIT@ models directly. Ultimately, we concluded that doing so would completely confound the influence of climate and habitat on species distribution. Given our

uncertainty in how climate and habitat interact to affect species distributions, we concluded that it was better to decouple these factors and treat climate separate from habitat. Thus, we developed a two-pronged modeling approach in which we model the species' climate-niche and habitat capability separately, and then combine the two along with prevalence to predict the distribution of the species. However, by maintaining separate climate and habitat components, it allows us to combine and intersect these in different ways to predict the future distribution of the species under alternative assumptions about how the species will respond to climate change. This seemed like the most credible approach for predicting future distributions of species.

5 Major Implementation Constraints

There are at least two major implementation constraints:

First, a constraint on building HABIT@ models (or any other type of model) is that the spatial data required to represent the landscape in a meaningful way may not exist for a species. Habit@ models (and all species distribution models) are only as good as our understanding and the data permit. For some species, we simply do not have the spatial environmental data needed to effectively model the species' habitat even if we have an ecological understanding of the species' habitat requirements. For such species, there is no solution other than to acquire the necessary spatial data, which is beyond the scope of this project. We attempted to account for this factor in the prioritization of representative species, whereby species with insufficient spatial data were assigned a low priority for modeling.

Second, the approach we developed to assess species distribution models relies entirely on extant presence/absence data. Readily available eBird data provide us with adequate data to assess the models for most bird species. Unfortunately, similar extant datasets do not exist for other taxonomic groups. Consequently, we will explore presence-only datasets that exist for herpetofauna (e.g., HerpNet, AmphibiaWeb) and will develop appropriate assessment methods as we encounter each unique situation.

6 Major Risks and Dependencies

The accuracy and utility of any species distribution model is completely dependent on the quality of the land cover and other environmental data described in Section 4.3.1. To date, a few shortcomings have been identified in the DSLland map that could have major implications on the quality of the models for forested species. For example, in the Pocomoke River pilot watershed the coastal plain hardwood forest ecological system is accurately mapped as a dominant native system in the watershed; however, a majority of this system has been subjected to extensive management, resulting in extensive conversion of this native system to loblolly pine plantation. This discrepancy would result in a significant underestimation of suitable habitat for pine-dependant species (e.g., brown-headed nuthatch) and a significant overestimation of suitable habitat for hardwood forest-dependant species (e.g., wood thrush). This deficiency has been remedied by mapping evergreen plantations (by the South Atlantic LCC group) in the NALCC. However, it is unclear what other spatial data deficiencies may exist, but this example points to the need

to carefully scrutinize the quality of all spatial data layers for their use in species distribution models.

Reliable, regional spatial data for forest vegetation characteristics and successional stage are notoriously difficult to acquire. The imputations of FIA data developed by the USFS Northern Research Station and described in the documentation on Disturbance and Succession, enable us to model a variety of forest vegetation attributes across the Northeast. However, there are several sources of inaccuracy of the FIA imputations as discussed in the Disturbance and Succession documentation that ultimately make these products unacceptable at the spatial resolution we are modeling habitat. Instead, we are using a much higher quality product of above-ground live biomass developed by USGS Woods Hole as modified by other data products on vegetation disturbance developed by the University of Maryland. But products of comparable quality are not yet available for other vegetation structure attributes that might be relevant to some species. Thus, our current species' distribution models are limited to the use of biomass to represent vegetation structure. For some species this may be entirely sufficient, but for others it may not.

It is important to acknowledge that 87 terrestrial species were selected to represent habitat clusters across the Northeast. Unfortunately, limited resources to date have allowed us to model only 30 species in phase 1 and 2 combined. Consequently, our current species-level ecological assessment is incomplete and will certainly change as more species are added to the mix. Thus, the results of our focal species assessment and the application of these results to landscape conservation design must be viewed with caution.

7 Literature Cited

- 2010 TIGER/Line Shapefiles Main Page. 2011. . Retrieved September 9, 2011, from <http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html>.
- Anderson, M. 2010. Creation of Regional Habitat Cover Maps: Application of the NE Terrestrial Habitat Classification System | rcngrants.org. Retrieved September 9, 2011, from <http://rcngrants.org/node/9>.
- Bailey, R. G., P. E. Avers, T. King, and W. H. McNab. 1994. Ecoregions and subregions of the United States (map). Washington, DC: US Geological Survey. Scale 1.
- Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *Forest Science* 49:12–35.
- Lawler, J. J., S. L. Shafer, D. White, P. Kareiva, E. P. Maurer, A. R. Blaustein, and P. J. Bartlein. 2009. Projected climate-induced faunal change in the Western Hemisphere. *Ecology* 90:588–597.
- Lobo, J. M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33:103–114.
- Matthews, S. N., L. R. Iverson, A. M. Prasad, and M. P. Peters. 2011. Changes in potential habitat of 147 North American breeding bird species in response to redistribution of trees and climate following predicted climate change. *Ecography*.

- McGarigal K, Compton BW, Plunkett EB, DeLuca WV, and Grand J. 2017. Designing sustainable landscapes products, including technical documentation and data products. https://scholarworks.umass.edu/designing_sustainable_landscapes/
- NALCC. 2011. Representative Species Executive Summary. Retrieved October 17, 2011, from http://www.fws.gov/northeast/science/pdf/Rep_Species_Executive_Summary.pdf.
- NatureServe Explorer: An Online Encyclopedia of Life. 2011. . Retrieved July 12, 2011, from <http://www.natureserve.org/explorer/>.
- Poole, A. 2005. The Birds of North America Online. Cornell Laboratory of Ornithology, Ithaca, NY.
- Poole, A. 2008. The birds of North America online. Cornell Lab.
- Ridgely, R. S., T. F. Allnutt, T. Brooks, D. K. McNicol, D. W. Mehlman, B. E. Young, and J. R. Zook. 2007. Digital Distribution Maps of the Birds of the Western Hemisphere, version 3.0. NatureServe. Arlington, VA, USA.
- Wiens, J. A., D. Stralberg, D. Jongsomjit, C. A. Howell, and M. A. Snyder. 2009. Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences* 106:19729.
- Wilson, B. T., Lister, A.J., Riemann, R.I., and Griffith, D.M. 2009. A moderate resolution nearest-neighbor imputation dataset of the continental United States for large area mapping of forest characteristics using field sampled data [database]. Northern Research Station, USDA Forest Service, St. Paul, Minnesota.
- Xian, G., C. Homer, and J. Fry. 2009. Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sensing of Environment* 113:1133–1147.

Appendix – Species Products

The following products are derived and distributed (from the DSL project website) for each modeled species. Note, there are many additional interim products stored internally but not distributed.

1. **speciesCN** - *Climate Niche (CN)* maps depicting each species' relative probability of climate suitability (0-1) based on the climate in 2010. Note, *Climate Niche Envelope (CNE)* binary maps depicting areas "inside" versus "outside" the species' *CNE* based on the climate in 2010 can be created by thresholding the *CN* grids at the specified cutpoint.
2. **speciesCN2080** - *Climate Niche (CN)* maps depicting each species' relative probability of climate suitability (0-1) based on the predicted climate in 2080 under the RCP 4.5 and 8.5 climate change scenarios (averaged across scenarios). Note, *Climate Niche Envelope (CNE)* binary maps depicting areas "inside" versus "outside" the species' *CNE* based on the predicted climate in 2080 under the RCP 4.5 and 8.5 climate change scenarios can be created by thresholding the *CN* grids at the specified cutpoint.
3. **speciesCZ2080** - *Climate Zones (CZ)* maps depicting zones of potential persistence, contraction and expansion in the future distribution of each species in 2080 under the RCP 4.5 and 8.5 climate change scenarios (averaged across scenarios).
4. **CNetable** - *Climate Niche (CN)* table summarizing each species' *CN* in 2010 and the proportional change in *CN* in 2030 and 2080 under the RCP 4.5 and 8.5 climate change scenarios (averaged across scenarios).
5. **speciesLC** - *Landscape Capability (LC)* maps depicting each species' relative probability of occurrence (0-1) based on the integrated climate-habitat-prevalence index in 2010.
6. **speciesLC2080** - *Landscape Capability (LC)* maps depicting each species' relative probability of occurrence (0-1) based on the integrated climate-habitat-prevalence index in 2080 under each landscape change simulation (averaged across urban growth and climate change simulations).
7. **speciesCR2080** - *Climate response* maps depicting the mean future *LC* calculated with current habitat and predicted future climate in 2080 (averaged across RCP4.5 and RCP8.5 scenarios) within the project area.
8. **speciesHR2080** - *Habitat response* maps depicting the mean future *LC* calculated with current climate and predicted future habitat in 2080 (averaged across urban growth simulations) within the project area.
9. **LCtable** - *Landscape Capability (LC)* table summarizing *LC* (in maximum-*LC* equivalent hectares) in 2010 and the proportional change in area in 2030 and 2080 (averaged across landscape change simulations) due to climate change, habitat change or both.